Enumerating 0-simple semigroups

Christopher Russell

University of
St Andrews

This dissertation is submitted in partial fulfilment for the degree of
Doctor of Philosophy (PhD)
at the University of St Andrews

October 2020

Declarations

Candidate's declaration

I, Christopher Russell, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 56,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in September 2016.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 30/04/21 Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date 04/05/21 Signature of supervisor

Permission for publication

InIn submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Christopher Russell, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date 30/04/21 Signature of candidate

Date 04/05/21 Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Christopher Russell, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date 30/04/21 Signature of candidate

Acknowledgements

I would like to acknowledge James Mitchell for being an excellent supervisor. I am lucky to have found a mentor whom I find it so easy to talk with. James was the one who introduced me to computational semigroup theory and during my doctoral studies he lit the initial spark for much of the research I undertook. I would also like to thank the Engineering and Physical Sciences Research Council (ESPRC) without whose funding I would not have been able to undertake this work.

Next, I would like to acknowledge my parents who have always supported me. They deserve a more attentive son, who visits more often. I am immensely grateful for all they have done for me and continue to do. I also want to thank their partners for taking an interest in me and showing me kindness.

Finally, I would like to acknowledge the many friends I have made during my eight years of study in St Andrews. Everything I have done academically feels inseparably linked to the many highs and occasional lows of my time here. All these experiences are the product of the people met along the way. There are the friends from my first year of undergraduate studies who continue to meet up like nothing has changed. There are the many generations of members from the university cross country club who now feel like a large widespread family. There are the other runners, coaches, colleagues, wardens, and housemates whom I am lucky to consider my friends. To acknowledge you all individually would require a whole chapter so this will have to do.

Abstract

Computational semigroup theory involves the study and implementation of algorithms to compute with semigroups. Efficiency is of central concern and often follows from the insight of semigroup theoretic results. In turn, computational methods allow for analysis of semigroups which can provide intuition leading to theoretical breakthroughs. More efficient algorithms allow for more cases to be computed and increases the potential for insight. In this way, research into computational semigroup theory and abstract semigroup theory forms a feedback loop with each benefiting the other.

In this thesis the primary focus will be on counting isomorphism classes of finite 0-simple semigroups. These semigroups are in some sense the building blocks of finite semigroups due to their correspondence with the Greens $\mathscr{D}$-classes of a semigroup. The theory of Rees 0-matrix semigroups links these semigroups to matrices with entries from 0-groups. Special consideration will be given to the enumeration of certain sub-cases, most prominently the case of congruence free semigroups. The author has implemented these enumeration techniques and applied them to count isomorphism classes of 0-simple semigroups and congruence free semigroups by order. Included in this thesis are tables of the number of 0-simple semigroups of orders less than or equal to 130, up to isomorphism. Also included are tables of the numbers of congruence free semigroups, up to isomorphism, with m Green's $\mathscr{L}$-classes and n Green's $\mathscr{R}$-classes for all $mn \leq 100$, as well as for various other values of m, n. Furthermore a database of finite 0-simple semigroups has been created and we detail how this was done. The implementation of these enumeration methods and the database are publicly available as **GAP** code. In order to achieve these results pertaining to finite 0-simple semigroups we invoke the theory of group actions and prove novel combinatorial results. Most notably, we have deduced formulae for enumerating the number of binary matrices with distinct rows and columns up to row and column permutations.

There are also two sections dedicated to covers of E-unitary inverse semigroups, and presentations of factorisable orthodox monoids, respectively. In the first, we explore the concept of a minimal E-unitary inverse cover, up to isomorphism, by defining various sensible orderings. We provide examples of Clifford semigroups showing that, in general, these orderings do not have a unique minimal element. Finally, we pose conjectures about the existence of unique

minimal E-unitary inverse covers of Clifford semigroups, when considered up to an equivalence weaker than isomorphism. In the latter section, we generalise the theory of presentations of factorisable inverse monoids to the more general setting of factorisable orthodox monoids. These topics were explored early in the authors doctoral studies but ultimately in less depth than the research on 0-simple semigroups.

Table of contents

List of figures

List of tables

Symbols

1_S Identity element of the monoid S 2

0_S Zero element of the semigroup with zero S 2

$E(S)$ The subset, sometimes the subsemigroup, of idempotents of S 2

$\vee$ The join operation 2

$\wedge$ The meet operation 2

$[s]_\rho$ The equivalence class of the element s with respect to the equivalence relation ρ 3

$[s]$ The equivalence class of the element s 3

S/ρ Quotient semigroup 3

Δ_S The trivial congruence of S 3

∇_S The universal congruence of S 3

$\operatorname{im} f$ The image of the function f 3

$\cong$ An equivalence relation on semigroups which relates those which are isomorphic 3

S^1 The semigroup S with identity adjoined 4

S^1x Principal left ideal of S generated by x 4

xS^1 Principal right ideal of S generated by x 4

S^1xS^1 Principal ideal$f$ of S generated by x 4

$\mathscr{L}$ Green's $\mathscr{L}$ relation 4

$\mathscr{R}$ Green's $\mathscr{R}$ relation 4

$\mathscr{H}$ Green's $\mathscr{H}$ relation 4

$\mathscr{D}$ Green's $\mathscr{D}$ relation 4

$\mathscr{J}$ Green's $\mathscr{J}$ relation 5

$\mathscr{M}^0[G;I,J;P]$ Rees zero matrix semigroup 5

$\mathbf{m}$ The set $\{1,2,\ldots,n\}$ 6

$\mathscr{M}^0[G;P]$ Rees zero matrix semigroup 6

x^g The image of x under the action of g 7

x^G The orbit of x with respect to the action of G 8

G_x The stabilizer of x with respect to the action of G 8

$\mathrm{fix}(g)$ The fix of an element g of a group G 8

$X//G$ The set of orbits of X with respect to a group action of G 8

$\mathscr{P}(X)$ The power set of X 9

$M_{m\times n}(G_0)$ The set of $m\times n$ matrices with entries from the 0-group G_0 15

Y^X The set of all functions from X to Y. 23

$\mathbf{1}$ The trivial group 26

$\mathrm{Inn}(G)$ The inner automorphism group of the group G 28

$\mathrm{Out}(G)$ The outer automorphism group of the group G 28

$\wr$ The wreath product 28

C_{p^k} The cyclic group of order p^k 41

$(\mathbb{Z}/p^k\mathbb{Z})^+$ The additive group of integers modulo p^k 41

$(\mathbb{Z}/p^k\mathbb{Z})^\times$ The multiplicative group of integers modulo p^k 41

P^T The transpose of the matrix P 57

$\mathscr{S}(P)$ The binary shape of the matrix P 58

$\mathbf{1}^0$ The trivial group with a zero element adjoined 58

$\equiv$ An equivalence relation on matrices which relates those which are equivalent up to row and column permutations 58

$M_{m\times n}(G^0,S)$ The set of $m\times n$ matrices with entries from the 0-group G_0 with binary shape equal to S 59

$B(T)$ The bipartite graph associated with the normal type T. 69

$\{0,1\}^{\mathbf{m}\times\mathbf{n}}$ The set of all functions from $\{0,1\}$ to $\mathbf{m}\times\mathbf{n}$, which represent $m\times n$ binary matrices 84

Chapter 1

Background

In this section we present some background theory requisite to most, if not all, of the remainder of this thesis. At the start of each chapter we may also present further background material relevant to that chapter only. All definitions can be found with the help of the index, and a description of any notation used can be found in the Symbols glossary. Chapter 4 has its own glossary of notation Chapter 4 Symbols. Our convention will be to write functions on the right. We will break this convention for cases such as the image of a function $\operatorname{im} f$, the kernel of a function $\ker f$, or the idempotents of a semigroup $E(S)$. In these cases it is standard practice to write these functions on the left.

1.1 Semigroups

Let X be a set and let $* : X \times X \to X$ be a binary operation on X. The operation $*$ is said to be *associative* if

$$(x * y) * z = x * (y * z)$$

holds for all x, y, z in X. A *semigroup* is a set together with an associative binary operation. We may denote the semigroup formed by the set X and the associative binary operation $*$ on X as the pair $(X, *)$. Often we will simply denote a semigroup by S in which case $s \in S$ indicates that s is an element of the underlying set of S and multiplication will be denoted by concatenation.

Amongst the elements of a semigroup there will be some special cases to which we repeatedly refer. An element e of a semigroup satisfying $ee = e$ is called an *idempotent*. The *zero element* $0 \in S$ denotes a special idempotent which satisfies $0s = s0 = 0$ for all s in S. The *identity element* $1 \in S$ is an idempotent which satisfies $1s = s1 = s$ for all s in S. Semigroups may have either: a zero, an identity, both, or neither. If present, then a zero or an identity is

unique. If these special elements exist in a semigroup S, then we may write 1_S and 0_S to refer to the identity and the zero of S, respectively. A semigroup with an identity is called a *monoid*. Idempotents are amongst the most important elements of a semigroup for gleaning information about the structure of the semigroup as a whole. We will denote the idempotents of a semigroup S by $E(S)$. In some cases $E(S)$ will be a subsemigroup of S in which case S is said to be an *orthodox* semigroup.

An element s of a semigroup S has an *inverse* if there exists $t \in S$ such that $sts = s$ and $tst = t$. An element may have no inverses, one inverse, or multiple inverses. A semigroup where every element has exactly one inverse is called an *inverse semigroup*. In this case, it is typical to denote the unique inverse of an element s by s^{-1}. More generally, a semigroup where every element has at least one inverse is called a *regular semigroup*.

1.2 Orders

Here we define several types of order relations and some important orders on semigroups. A relation on a set X is a subset of $X \times X$. Let X be a set and let $\rho \subseteq X \times X$. Then:

(i) ρ is a *reflexive relation* if $(x,x) \in \rho$ for all $x \in X$;

(ii) ρ is a *symmetric relation* if $(x,y) \in \rho$ implies $(y,x) \in \rho$ for all $x,y \in X$;

(iii) ρ is a *anti-symmetric relation* if $(x,y),(y,x) \in \rho$ implies $x = y$ for all $x,y \in X$;

(iv) ρ is a *transitive relation* if $(x,y),(y,z) \in \rho$ implies $(x,z) \in \rho$ for all $x,y,z \in X$;

(v) ρ is a *connex relation* if $x,y \in X$ implies at least one of: $(x,y) \in \rho$, and $(y,x) \in \rho$.

These properties are called reflexivity, symmetry, anti-symmetry, transitivity, and connexity. We note that connexity implies reflexivity. A *partial order* is a reflexive, anti-symmetric and transitive relation. The *join* of two elements x_1, x_2 of a partial order $(X, \leq)$ is their least upper bound and is denoted by $x_1 \vee x_2$. More generally, if Y is a finite subset of the partial order $(X, \leq)$ then the join $\vee Y$ is the set of elements x of X such that: $x \geq y$ for all $y \in Y$, and if $z \in X$ such that $z \geq y$ for all $y \in Y$ then $z \geq x$. A *join semilattice* is a partial order such that every finite non-empty subset has a unique least upper bound. Similarly, the *meet* of two elements x_1, x_2 of a partial order is their greatest lower bound. A *meet semilattice* is a partial order which has a unique greatest lower bound for any finite non-empty subset. The meet operation is denote by $\wedge$ and is used in the same manner as the join operation. A *lattice* is a partial order which is both a meet-semilattice and a join-semilattice. A *total order* is a relation which is anti-symmetric, transitive and connex. A total order is a special type of lattice. A *well order* is a total order such that every non-empty subset has a least element.

1.3 Congruences

Congruences are the semigroup theoretic method of handling homomorphisms and we will give a brief overview in this section. Throughout this thesis we will use square bracket notation to denote equivalence classes. If ρ is an equivalence on a set S then $[s]_\rho$ denotes the equivalence class of an element s of S and we may also merely write $[s]$ when ρ is clear from the context. An equivalence relation ρ on a semigroup S is called a *congruence* if and only if

$$(sx, ty) \in \rho$$

is satisfied for all pairs $(s,t),(x,y)$ in ρ. If ρ is a congruence on S then S/ρ denotes the quotient semigroup which has element set $\{[s]_\rho : s \in S\}$, and multiplication defined by $[s][t] = [st]$. It follows from the definition of a congruence that this multiplication is well defined. For any semigroup S the *trivial congruence* refers to the relation

$$\Delta_S = \{(s,s) : s \in S\},$$

and the *universal congruence* refers to the relation

$$\nabla_S = \{(s,t) : s,t \in S\}.$$

The relations Δ_S and ∇_S are always congruences and their quotients are isomorphic to S and the trivial semigroup, respectively. The first isomorphism theorem for semigroups states that if $f : S \to T$ is a semigroup homomorphism then the image $\operatorname{im} f$ is a subsemigroup of T, the relation given by $(s,s') \in \psi$ if and only if $f(s) = f(s')$ is a congruence on S, and the semigroups S/ψ and $\operatorname{im} f$ are isomorphic. In this thesis we will write $S \cong T$ to denote that the semigroups S and T are isomorphic.

1.4 Ideals

Ideals are special types of subsemigroups and they are relevant to certain congruences, and Green's relations (which will be discussed in Section 1.5). Let A be subset of a the semigroup S. Then A is:

(i) a *left ideal* if $SA \subseteq A$;

(ii) a *right ideal* if $AS \subseteq A$; and

(iii) a *(two-sided) ideal* if it is both a left ideal and a right ideal.

All left, right, and two-sided ideals are subsemigroups but the converse need not be true. Note that when we refer to a subsemigroup as an 'ideal' then this always means a two-sided ideal. A (left, right, or two-sided) ideal of S is called *proper* if it is a proper subset of S which is not equal to $\{0_S\}$ in the case that S has a zero element. The smallest (left, right, or two-sided) ideal containing a set X is called the (left, right, or two-sided) ideal generated by X. The special case of a (left, right, or two-sided) ideal generated by a singleton $\{x\}$ is called *principal (left, right, or two-sided) ideal*. Let S be a semigroup and define

$$S^1 = \begin{cases} S & S \text{ has an identity element} \\ S \cup \{1\} & S \text{ has no identity element.} \end{cases}$$

where, in the latter case, $1s = s1 = s$ for all $s \in S^1$ and otherwise multiplication is the same as in S. This process is called adjoining an identity. It is convenient to denote left, right, and two-sided principal ideals generated by the element x of S by S^1x, xS^1, and S^1xS^1 (respectively).

The relation between ideals and congruences is as follows. Let I be a proper ideal of a semigroup S. Then the relation

$$\rho_I = (I \times I) \cup \{(s,s) : s \in S\}$$

is a congruence on S called a *Rees congruence*. Unlike in ring theory, there need not be a bijective correspondence between the ideals of a semigroup and its congruences.

1.5 Green's relations

In this section we define Green's relations, which were first described by Green in 1951 [12]. Green's relations contain a significant amount of information about a semigroup and are key tools in understanding the structure of a semigroup. Their importance can hardly be made more apparent than by noting that the distinguished semigroup theorist John M. Howie described these relations as "so all-pervading that, on encountering a new semigroup, almost the first question one asks is 'What are the Green's relations like?' " [20]. We now define Green's relations:

(i) $s \mathscr{L} t$ if and only if $S^1 s = S^1 t$;

(ii) $s \mathscr{R} t$ if and only if $sS^1 = tS^1$;

(iii) $s \mathscr{H} t$ if and only if $s \mathscr{L} t$ and $s \mathscr{R} t$;

(iv) $s \mathscr{D} t$ if and only if there exists some u in S such that $s \mathscr{L} u \mathscr{R} t$;

(v) $s \mathscr{J} t$ if and only if $S^1sS^1 = S^1tS^1$;

for all $s,t \in S$. We note that all five of these relations are equivalences [19, §2.1] and that $\mathscr{D} = \mathscr{J}$ if S is finite [19, §2.1]. In this thesis we are primarily concerned with finite semigroups and so $\mathscr{J}$ will often be omitted from our considerations.

1.6 0-simple semigroups

A semigroup is called *simple* if it has no proper ideals. A semigroup S with zero is called 0-simple if:

(i) S has no proper ideals except $\{0_S\}$, and

(ii) $\{st : s,t \in S\} \neq \{0_S\}$.

Condition (ii) only exists to exclude the case where S is the two element null semigroup (a semigroup where all products are equal to 0). The importance of 0-simple semigroups follows from a decomposition theorem for such semigroups. Briefly, if S is a semigroup, $a \in S$ is not equal to 0_S, and J_a is the $\mathscr{J}$-class of S containing a, then the quotient of the principal ideal S^1aS^1 by the semigroup containing all elements of S^1aS^1 except those in J_a is a 0-simple semigroup. In finite semigroups $\mathscr{D} = \mathscr{J}$ so in this case the decomposition says that the ideal generated by a $\mathscr{D}$-class (except the $\mathscr{D}$-class of the zero element) behaves like a 0-simple semigroup when we identify all the elements not in that $\mathscr{D}$-class with the zero element. It is important to realise that 0-simple semigroups are not the analogue of simple groups in the sense of having no proper homomorphic images. Semigroups with no proper homomorphic images are called congruence free semigroups and are discussed further in Section 1.7.

A key construction in the study of 0-simple semigroups is the Rees 0-matrix semigroup. Let G be a group. Let I,J be non-empty sets. Let $P = (p_{i,j})$ be a matrix with entries indexed by $I \times J$ and with entries which are elements of the 0-group $G^0 = G \cup \{0\}$. We will say that P is *regular* if it has no rows consisting entirely of 0 and no columns consisting entirely of 0. Then we may define the *Rees 0-matrix semigroup* to be the set $(I \times G \times J) \cup \{0\}$ with multiplication defined by

$$(i_1,g,j_1)(i_2,h,j_2) = \begin{cases} (i_1, gp_{i_2,j_1}h, j_2) & p_{i_2,j_1} \neq 0, \\ 0 & p_{i_2,j_1} = 0, \end{cases}$$

$$(i_1,g,j_1)0 = 0(i_1,g,j_1) = 00 = 0.$$

This semigroup is typically denoted by $\mathscr{M}^0[G;I,J;P]$.

Rees showed that every completely 0-simple semigroup is isomorphic to some Rees 0-matrix semigroup [36]. Being completely 0-simple is equivalent to being 0-simple for finite semigroups. As our focus will be finite 0-simple semigroups we will not define completely 0-simple semigroups nor will we present the theory of Rees 0-matrix semigroups beyond what is necessary for finite 0-simple semigroups. We now present a version of the aforementioned result due to Rees which only covers the case of finite 0-simple semigroups.

Theorem 1.6.1. *Let G^0 be a finite 0-group, let I,J be finite non-empty sets, and let $P=(p_{i,j})$ be a matrix indexed by $I\times J$ with entries in G^0. Furthermore suppose that P is regular. Then $\mathscr{M}^0[G;I,J;P)]$ is a finite 0-simple semigroup.*

Conversely, every finite 0-simple semigroup is isomorphic to some Rees 0-matrix semigroup constructed in this way.

Furthermore, the precise conditions for two Rees 0-matrix semigroups to be isomorphic are known. This result will be fundamental to our efforts to enumerate 0-simple semigroups up to isomorphism in Chapter 2.

Theorem 1.6.2 ([19, Theorem 3.4.1])**.** *Let $S=\mathscr{M}_0(G;I,J;P)$ and $T=\mathscr{M}_0(K;J,M;Q)$ be Rees 0-matrix semigroups. Then S and T are isomorphic if and only if there exists an isomorphism $\theta:G\to K$, bijections $\rho:I\to J$, $\sigma:J\to M$, and elements $f_i\in K$ for all $i\in I\cup J$ such that*

$$p_{i,j}\theta = f_i^{-1}(q_{i\rho,j\sigma})f_j$$

for all $i\in I$ and $j\in J$.

This theorem does well to describe the situation in full generality. However, for the goal of enumerating isomorphism classes there are a few conventions we will introduce to simplify notation without loss of generality. For a Rees 0-matrix semigroup $\mathscr{M}_0(G;I,J;P)$ the cardinality of the index sets I and J are important but the specific elements are not. Let $m=|I|$ and $n=|J|$. In the following chapters of this thesis m,n will always be finite and it will be simpler to have I and J always be the sets $\{1,\dots,m\}$ and $\{1,\dots,n\}$, respectively. We will denote the set $\{1,\dots,m\}$ by $\mathbf{m}$ throughout this thesis. We will often abbreviate $\mathscr{M}^0[G;\mathbf{m},\mathbf{n};P]$ to $\mathscr{M}^0[G;P]$ as the m,n are known from the dimensions of P.

1.7 Congruence free semigroups

A semigroup is said to be *congruence free* if the only congruences on that semigroup are the trivial congruence and the universal congruence. Thus, these semigroups are arguably the

semigroup analogue of simple groups. The following theorem classifies these semigroups into three categories.

Theorem 1.7.1 ([19, Theorem 3.7.1, Theorem 3.7.2])**.** *Let S be a finite semigroup. Then S is congruence free if and only if one of the following holds:*

(i) *S is a simple group.*

(ii) *S has order less than or equal to 2.*

(iii) *S is isomorphic to a Rees 0-matrix semigroup $\mathscr{M}^0[G;I,J;P]$ where G is the trivial group, and P is regular with all rows distinct and all columns distinct.*

The enumeration of finite congruence free semigroups of type (iii) is a novel area of research which we undertake in Chapter 4.

1.8 Group actions

Let G be a group and let X be a set. A *left group action* is a function $f : G \times X \to X$ satisfying:

$$f(1_G, x) = x$$

$$f(gh, x) = f(g, f(h, x))$$

for all $g, h \in G$ and $x \in X$. We will often denote $f(g,x)$ by the more succinct gx when f is clear from context. A *right group action* is a function $f : X \times G \to X$ satisfying:

$$f(x, 1_G) = x$$

$$f(x, gh) = f(f(g,x), h)$$

for all $g, h \in G$ and $x \in X$. We will often denote $f(x,g)$ using the superscript notation x^g when f is clear from context. Whether left or right, we will say that the group G acts on the set X.

We will use right actions for the remainder of this section and note that the situation for left actions is largely analogous. The group action $f : G \times X \to X$ can be understood as a group homomorphism from the acting group G into the automorphism group of X. To be explicit, this homomorphism sends $g \in G$ to the following automorphism of X:

$$x \mapsto f(g,x).$$

We call this homomorphism the *representation* of the corresponding group action. The term *representation* is also used to refer the image of such a homomorphism.

The set of images of $x \in X$ under the action of all elements of G is called the *orbit* of x, with respect to the action of G, and is denoted by:

$$x^G = \{x^g : g \in G\}.$$

The *stabilizer* of $x \in X$ is the subset of G which fixes x, and is denoted by:

$$G_x = \{g \in G : x^g = x\}.$$

This stabilizer is necessarily a subgroup of G. There is a close link between the orbits and the stabilizers of a group action. This is described in the orbit-stabilizer theorem which states: the size of the orbit x^G is equal to the index of the stabilizer group $|G : G_x|$. It is sometimes useful to refer to the *fix* of an element g of G which is equal to

$$\mathrm{fix}(g) = \{x \in X : x^g = x\},$$

in other words, the set of elements of X fixed by G.

The *kernel* of the action of G on X is the following subgroup of G:

$$\{g \in G : x^g = x \text{ for all } x \in X\}.$$

This subgroup is the same as the kernel of the homomorphism which is called the representation of G. A group action is said to be *faithful* if and only if the kernel is of the action is trivial. The kernel must be a normal subgroup of G and, where the action of G is not faithful, a faithful action may be created by taking the quotient of G by the kernel. Computationally, it is often more efficient to work with smaller group actions.

In Chapters 2 and 4 group actions will play a key role and we will want to determine the number of orbits of specific actions. We will denote the collection of all orbits of X with respect to the action of G by $X//G$. The following result, known as both the orbit-counting theorem and Burnside's lemma, is a well known way to count orbits.

Theorem 1.8.1 (Orbit-counting theorem)**.** *Let X be a set and let G be a group which acts on X. Then*

$$|X//G| = \frac{1}{|G|} \sum_{g \in G} |X^g|$$

where X^g denotes the elements in X fixed by g.

Given a group action involving a group G and a set X there is a natural way to define an action of G on the power set $\mathscr{P}(X)$ of X. This is known as an *induced action* of G on $\mathscr{P}(X)$ and is defined by

$$Y^g = \{y^g : y \in Y\} \tag{1.1}$$

for all $Y \subset X$ and $g \in G$. Essentially G acts on a subset of X by action on every element of that subset simultaneously. There may also be proper subsets Q of $\mathscr{P}(X)$ for which Equation 1.1 defines an action. This can happen as long as Q is 'closed' under the induced action of G, that is to say if Y is a subset of X contained in Q then Y^g is in Q for all g in G. We will call any action of G on subsets of X defined by Equation 1.1 an induced action, not just the action of G on the power set of X.

1.9 Graphs and digraphs

A *directed graph*, or *digraph*, is a set of vertices together with a set of directed edges. The edge set of a digraph with vertex set X is a subset of $X \times X$ where the element (x,y) can be interpreted as an edge from x to y. Let G be a digraph, then it is common to refer to the vertex set of G as $V(G)$ and the edge set as $E(G)$. A digraph G may be denoted by the pair $(V(G),E(G))$ of the vertex set and edge set. Often we will refer to a vertex v or an edge (u,v) as being 'in' the digraph G, when technically it would be correct to say that v is in $V(G)$ or (u,v) is in $E(G)$.

The *adjacency matrix* of a digraph G is a binary matrix $(p_{u,v})_{u,v\in V(G)}$ indexed by $V(G) \times V(G)$ such that $p_{u,v} = 1$ if and only if (u,v) is an edge of G. A *path* on a digraph G is a sequence of distinct vertices $(v_1,v_2,\ldots,v_n)$ such that (v_i,v_{i+1}) is an edge of G for all $1 \leq i < n$. A digraph is called *(strongly) connected* if for every pair of vertices u,v there is a path which starts at u and ends at v. A digraph which is not connected is called *disconnected*. The *(strongly) connected components* of a digraph are the maximal connected subdigraphs. Each vertex and edge belongs to exactly one connected component.

Chapter 2

Counting 0-simple semigroups

Rees showed that all completely 0-simple semigroups are isomorphic to some Rees 0-matrix semigroup [36]. This result reveals a correspondence between these semigroups and matrices with entries from a 0-group[1] . The combinatorial nature of matrices together with results pertaining to isomorphisms tempt us to enumerate isomorphism classes of these semigroups.

It is an immediate consequence of Theorem 1.6.2 that if two Rees 0-matrix semigroups are isomorphic then the associated groups are isomorphic and the index sets of the associated matrices have the same cardinality. It then makes sense to separate the enumeration of Rees 0-matrix semigroups isomorphism classes into cases based on these three parameters: the isomorphism classes of the associated groups, the number of rows, and the number of columns. For each case counting isomorphism classes corresponds to counting orbits of certain group actions on matrices. We will apply Pólya enumeration theory to count these orbits. Our aim is to implement an efficient method for computing these counts.

In 1978 Houghton [17] noted that counting isomorphism classes of 0-simple semigroups constructed from $m \times n$ matrices over the 0-group G^0 can be achieved by counting orbits of matrices with respect to the split extension of G^{m+n} by $S_m \times S_n \times \mathrm{Aut}(G)$. He then applied Burnside's orbit counting lemma to deduce a formula for the number of isomorphism classes. We now give an overview of this formula. Let $\theta(j,k)$ denote an element of $S_m \times S_n$ with cycle type[2] $(j_1,\ldots,j_m)$ for the S_m part and $(k_1,\ldots,k_n)$ for the S_n part. Furthermore let ω be an element of $\mathrm{Aut}(G)$, the automorphism group of G, and let t be an element of G^{m+n}. Houghton writes $F(\theta(j,k),\omega,t)$ to denote the number of matrices fixed by $(\theta(j,k),\omega,t)$. Then Houghton's formula is essentially the following:

[1] A zero group is a group with a zero element adjoined.

[2] A element of S_m which has cycle type $(j_1,\ldots,j_m)$ is a permutation which has precisely j_1 1-cycles, j_2 2-cycles, ..., and j_m m-cycles. Each conjugacy class of the symmetric group consists precisely of all the elements of a specific cycle type.

$$\frac{1}{|G|^{m+n}|\operatorname{Aut}(G)|} \sum_{\substack{j_1+\cdots+mj_m=m \\ k_1+\cdots+nk_n=n \\ j_1,\ldots,j_m,k_1,\ldots,k_n\in\mathbb{N}}} F(\theta(j,k),\omega,t)/\left(\prod_{e=1}^{m}(j_e!e^{j_e})\prod_{f=1}^{n}(k_f!f^{k_f})\right)$$

Note that the sum is over all possible cycle types $(j_1,\ldots,j_m)$ and $(k_1,\ldots,k_m)$ of the S_m and S_n parts of $\theta(j,k)$, respectively. Also note that the number of elements of S_m with cycle type $(j_1,\ldots,j_m)$ is equal to $m!/\prod_{e=1}^{m}(j_e!e^{j_e})$. Houghton goes on to note that $F(\theta(j,k),\omega,t)$ is equal to the number of solutions $(g_0,\ldots,g_{d-1})\in(G^0)^d$ of a collection of equations. These equations describe the conditions on the entries of a matrix necessary for it to be fixed by the action of $(\theta(i,j),\omega,t)$. However there is no clear path from these equations to a combinatorial formula for the number of solutions which can be evaluated.

Further research was done by Houghton in counting the isomorphism classes of 0-simple semigroups over elementary abelian groups [18]. Whilst the author was undertaking this research he was initially unaware of Houghton's work.

2.1 Isomorphisms

Theorem 1.6.2 does well to describe the isomorphism situation for Rees 0-matrix semigroups in full generality. It is worth noting that this theorem does not provide an effective blueprint for showing any two given Rees 0-matrix semigroups are isomorphic. It tells us that two Rees 0-matrix semigroups are isomorphic if suitable ρ,σ,f_1,f_2, and θ exist but does not give us a method for determining if they exist nor a method for constructing them when they do exist. In this section we will specialize the setting of Theorem 1.6.2 to one simpler but sufficient for our goals. We will also describe the implications of this theorem in more detail in hopes of increasing the reader's intuition regarding isomorphisms between Rees 0-matrix semigroups.

For the goal of enumerating isomorphism classes it will be helpful to narrow our focus and only consider the variables which determine the isomorphism class of a 0-simple semigroup. For a Rees 0-matrix semigroup $\mathscr{M}_0(G;I,J;P)$ the cardinality of the index sets I and J affect the isomorphism class but the specific elements in these sets do not. Similarly, the isomorphism class of G is important but the representation is not. The following corollary summarises these observations.

Corollary 2.1.1. *Let $S=\mathscr{M}_0(G;I,J;P)$ and $T=\mathscr{M}_0(H;I',J';Q)$ be Rees 0-matrix semigroups. If S is isomorphic to T then G is isomorphic to H, $|I|=|I'|$, and $|J|=|J'|$.*

The key consequence of Corollary 2.1.1 is that in order to consider all isomorphism classes of 0-simple semigroups we need only consider one representative of each isomorphism class of

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{1\rho^{-1},1} & a_{1\rho^{-1},2} & \dots & a_{1\rho^{-1},n} \\ a_{2\rho^{-1},1} & a_{2\rho^{-1},2} & \dots & a_{2\rho^{-1},n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m\rho^{-1},1} & a_{m\rho^{-1},2} & \dots & a_{m\rho^{-1},n} \end{pmatrix}$$

Apply $\rho \in S_m$ to rows

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{1,1\sigma^{-1}} & a_{1,2\sigma^{-1}} & \dots & a_{1,n\sigma^{-1}} \\ a_{2,1\sigma^{-1}} & a_{2,2\sigma^{-1}} & \dots & a_{2,n\sigma^{-1}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1\sigma^{-1}} & a_{m,2\sigma^{-1}} & \dots & a_{m,n\sigma^{-1}} \end{pmatrix}$$

Apply $\sigma \in S_n$ to the columns

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \longrightarrow \begin{pmatrix} u_1^{-1} * a_{1,1} & u_1^{-1} * a_{1,2} & \dots & u_1^{-1} * a_{1,n} \\ u_2^{-1} * a_{2,1} & u_2^{-1} * a_{2,2} & \dots & u_2^{-1} * a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ u_m^{-1} * a_{m,1} & u_m^{-1} * a_{m,2} & \dots & u_m^{-1} * a_{m,n} \end{pmatrix}$$

Multiply rows on the left by the inverses of $u_1, u_2, \dots, u_m \in G$, respectively.

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{1,1} * v_1 & a_{1,2} * v_2 & \dots & a_{1,n} * v_n \\ a_{2,1} * v_1 & a_{2,2} * v_2 & \dots & a_{2,n} * v_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} * v_1 & a_{m,2} * v_2 & \dots & a_{m,n} * v_n \end{pmatrix}$$

Multiply columns on the right by $v_1, v_2, \dots, v_m \in G$, respectively.

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \longrightarrow \begin{pmatrix} \theta(a_{1,1}) & \theta(a_{1,2}) & \dots & \theta(a_{1,n}) \\ \theta(a_{2,1}) & \theta(a_{2,2}) & \dots & \theta(a_{2,n}) \\ \vdots & \vdots & \ddots & \vdots \\ \theta(a_{m,1}) & \theta(a_{m,2}) & \dots & \theta(a_{m,n}) \end{pmatrix}$$

Apply the automorphism $\theta \in \mathrm{Aut}(G^0)$ to all entries.

Fig. 2.1 The ways which a matrix can be modified which do not change the isomorphism class of the related Rees 0-matrix semigroup.

groups, and the choice of the elements in the index sets is also irrelevant. Let G be a group and let $m,n>0$ be integers. Then we will say that a Rees 0-matrix semigroup $\mathscr{M}_0(G';I,J;P)$ is of type (G,m,n) if and only if G is isomorphic to G', $|I|=m$ and $|J|=n$. Going forward we can tackle these cases independently. Henceforth we may assume, without loss of generality, that the index sets are of the form $I=\mathbf{m}$ and $J=\mathbf{n}$. This will allow us to make statements with less notation. For brevity we will denote the Rees 0-matrix semigroup $\mathscr{M}_0(G;\mathbf{m},\mathbf{n};P)$ by $\mathscr{M}_0(G;P)$ since m and n are implicit from the dimensions of the matrix. The following corollary is a specialisation of Theorem 1.6.2 to the situation where the two Rees 0-matrix semigroups are over the same group, and formed from matrices with the same dimensions.

Corollary 2.1.2. *Let $S=\mathscr{M}_0(G;P)$ and $T=\mathscr{M}_0(G;Q)$ be Rees 0-matrix semigroups. Then $S\cong T$ if and only if there exists $\rho\in S_m$, $\sigma\in S_n$, $f_1\in G^m$, $f_2\in G^n$, and $\theta\in\mathrm{Aut}(G^0)$ such that*

$$q_{i,j}\theta=(if_1)^{-1}(p_{i\rho,j\sigma})(jf_2)$$

for all $i\in\mathbf{m}$ and $j\in\mathbf{n}$.

A good way to get a grasp on this result is to consider how we can change a given matrix and obtain another matrix which forms a Rees 0-matrix semigroup in the same isomorphism class. Let $P=(p_{i,j})$ be a $m\times n$ matrix with entries from the 0-group G^0. Then there are five essential ways in which we may modify P without changing the isomorphism class. These are as follows:

(i) Apply any permutation $\rho\in S_m$ to the m rows of P;

(ii) Apply any permutation $\sigma\in S_n$ to the n columns of P;

(iii) Choose any tuple $f_1=(g_1,\ldots,g_m)\in G^m$ of m elements from G and left multiply row i of P by g_i for all $i\in\mathbf{m}$;

(iv) Choose any tuple $f_2=(g_1,\ldots,g_n)\in G^n$ of n elements from G and left multiply column i of P by g_i^{-1} for all $i\in\mathbf{n}$;

(v) Pick an automorphism of G^0 and apply it to every entry of P.

See Figure 2.1 for an illustration of these five matrix modifications. Moreover, we may apply any combination of these modifications in any order and the isomorphism class of the altered matrix will not change[3].

[3]Note that a valid alternative is to left multiply the columns and right multiply the rows. However, in general, if you are combining a sequence of these manipulations then you cannot guarantee the isomorphism class will be fixed if you were to both left multiply the rows and the columns, or right multiply both.

2.2 Group actions on matrices over 0-groups

Let us fix a group G and integers $m,n > 0$. We will denote by $M_{m\times n}(G_0)$ the set all $m \times n$ matrices with entries from G^0. For any of the methods of modifying a matrix described at the end of Section 2.1 the function which applies that modification to the elements of $M_{m\times n}(G^0)$ is in fact a permutation of $M_{m\times n}(G^0)$. Recall that the representation of a group action may refer to the related homomorphism from the group into the symmetric group of the set which it acts upon or the image of that homomorphism. The action of the combination of all permutations on the rows and columns of matrices in $M_{m\times n}(G^0)$ has representation isomorphic to $S_m \times S_n$. The collection of all tuples of length m and n of elements of G acting on $M_{m\times n}(G^0)$ by the previously described mix of row-wise and column-wise multiplication form a group with representation isomorphic to G^{m+n}, the direct product of $m+n$ copies of G. The automorphisms of G^0 applied entry-wise to matrices in $M_{m\times n}(G^0)$ generates a group with representation isomorphic to $\mathrm{Aut}(G^0)$, or equivalently $\mathrm{Aut}(G)$[4]. Roughly speaking, the group generated by the combination of the actions of $S_m \times S_n$, G^{m+n}, and $\mathrm{Aut}(G^0)$ has orbits corresponding to isomorphism classes of Rees 0-matrix semigroups of type (G,m,n). In this section we define this group action and an alternative representation of the elements of $M_{m\times n}(G^0)$ which will allow for more efficient enumeration of isomorphism classes.

Recall that when a group G acts on a set X there is a homomorphism $G \to S_X$ such that $x^g = g(x)$ for all $x \in X$. The image of this homomorphism is called the representation of the action. We are considering five actions:

$$S_m \ni \rho : (p_{i,j}) \mapsto (p_{i\rho,j}),$$

$$S_n \ni \sigma : (p_{i,j}) \mapsto (p_{i,j\sigma}),$$

$$G^m \ni (g_1,\ldots,g_m) : (p_{i,j}) \mapsto (g_i^{-1}p_{i,j}),$$

$$G^n \ni (h_1,\ldots,h_n) : (p_{i,j}) \mapsto (p_{i,j}h_j), \text{ and}$$

$$\mathrm{Aut}(G^0) \ni \theta : (p_{i,j}) \mapsto (p_{i,j}\theta).$$

We now define an action which has representation isomorphic to the group generated by the representations of these actions. That group will be a semidirect product of G^{m+n} by $S_m \times S_n \times \mathrm{Aut}(G^0)$. We want to create a group with a set of elements equal to the Cartesian product of G^m, G^n, S_m, S_n and $\mathrm{Aut}(G^0)$. It is easy to see[5] that the actions of S_m, S_n, and

[4] Note any automorphism of G^0 must fix the zero element so these two automorphism groups are isomorphic.

[5] Consider left multiplying the entries of the first row of a matrix in $M_{m\times n}(G^0)$ by some element $g \in G$. Generally, if we were to also apply the permutation $(1,2)$ to swap rows 1 and 2 then the result will be different if we do this before or after the aforementioned row multiplication. Also, if $\theta \in \mathrm{Aut}(G^0)$ is such that $g\theta \neq g$ then

$\mathrm{Aut}(G^0)$ commute and that the action of G^m and G^n need not commute with any of the other three actions. Since these actions do not necessarily commute the group we create is not the direct product of these five groups. However, it turns out that the expected behaviour can be captured in a semidirect product. Let $g = (g_1, \dots g_m), h = (h_1, \dots, h_m) \in G^m; u = (u_1, \dots, u_n), v = (v_1, \dots, v_n) \in G^n; \rho_1, \rho_2 \in S_m; \sigma_1, \sigma_2 \in S_n$; and $\theta_1, \theta_2 \in \mathrm{Aut}(G^0)$. Then we want to define the action

$$(p_{i,j}) \cdot (g, u, \rho_1, \sigma_1, \theta_1) = (g_i^{-1} p_{i\rho_1, j\sigma_1} \theta_1 u_j) \tag{2.1}$$

together with a multiplication $*$ compatible with this action, i.e. the following equation must hold:

$$((p_{i,j}) \cdot (g, u, \rho_1, \sigma_1, \theta_1)) \cdot (h, v, \rho_2, \sigma_2, \theta_2) = (p_{i,j}) \cdot ((g, u, \rho_1, \sigma_1, \theta_1) * (h, v, \rho_2, \sigma_2, \theta_2)).$$

To determine how to define this multiplication we evaluate the left hand side.

$$\begin{aligned}
&((p_{i,j}) \cdot (g, u, \rho_1, \sigma_1, \theta_1)) \cdot (h, v, \rho_2, \sigma_2, \theta_2) \\
=&(g_i^{-1} p_{i\rho_1, j\sigma_1} \theta_1 u_j) \cdot (h, v, \rho_2, \sigma_2, \theta_2) \\
=&(h_i^{-1} (g_{i\rho_2}^{-1} p_{i\rho_2\rho_1, j\sigma_2\sigma_1} \theta_1 u_{j\sigma_2}) \theta_2 v_j) \\
=&h_i^{-1} g_{i\rho_2}^{-1} \theta_2 \cdot p_{i\rho_2\rho_1, j\sigma_2\sigma_1} \theta_1\theta_2 \cdot u_{j\sigma_2} \theta_2 v_j.
\end{aligned}$$

Therefore, in order for our intended action to work, we must define

$$(g, u, \rho_1, \sigma_1, \theta_1) * (h, v, \rho_2, \sigma_2, \theta_2)$$

to be equal to

$$((g_{1\rho_2} \theta_2 h_1, \dots, g_{m\rho_2} \theta_2 h_m), (u_{1\sigma_2} \theta_2 v_1, \dots, u_{n\sigma_2} \theta_2 v_n), \rho_2\rho_1, \sigma_2\sigma_1, \theta_1\theta_2). \tag{2.2}$$

In other words we can define the semidirect product $(G^m \times G^n) \rtimes_\psi (S_m \times S_n \times \mathrm{Aut}(G^0))$ where the homomorphism $\psi : S_m \times S_n \times \mathrm{Aut}(G^0) \to \mathrm{Aut}(G^m \times G^n)$ is defined by

$$(\rho, \sigma, \theta)\psi = ((g_1, \dots, g_m), (u_1, \dots, u_n)) \mapsto ((g_{1\rho}\theta, \dots, g_{m\rho}\theta), (u_{1\sigma}\theta, \dots, u_{n\sigma}\theta)). \tag{2.3}$$

We now prove this is a group which acts by Equation 2.1 on the set $M_{m\times n}(G^0)$.

applying θ to all entries of the matrix before or after the aforementioned row multiplication produces a different result. This shows how the action of G^m does not commute with the actions of S_m, S_n, and $\mathrm{Aut}(G^0)$. The same is true for G^n and can be shown by an analogous example.

Lemma 2.2.1. *Let G be a finite group, and let $m,n > 0$ be integers. Then the set $G^m \times G^n \times S_m \times S_n \times \mathrm{Aut}(G^0)$ together with the multiplication $*$ defined in Equation 2.2 is a group. Furthermore this group acts on the set $M_{m\times n}(G^0)$ by*

$$(p_{i,j}) \cdot (g,u,\rho,\sigma,\theta) = (g_i^{-1} p_{i\rho,j\sigma}\theta u_j).$$

Proof. The element $(1_{G^m}, 1_{G^n}, 1_{S_m}, 1_{S_n}, 1_{\mathrm{Aut}(G^0)})$ composed of the identity elements of the factors is the identity with respect to $*$. Let $g = (g_1,\ldots,g_m) \in G^m, u = (u_1,\ldots,u_n) \in G^n, \rho \in S_m, \sigma \in S_n$, and $\theta \in \mathrm{Aut}(G^0)$. Then the inverse of the element (g,u,ρ,σ,θ) is

$$((g_{1\rho^{-1}}^{-1},\theta^{-1},\ldots,g_{m\rho^{-1}}^{-1}\theta^{-1}),(u_{1\sigma^{-1}}^{-1}\theta^{-1},\ldots,u_{n\sigma^{-1}}^{-1}\theta^{-1}),\rho^{-1},\sigma^{-1}\theta^{-1}).$$

Finally we prove associativity. Let

$$g = (g_1,\ldots g_m), h = (h_1,\ldots,h_m), k = (k_1,\ldots,k_m) \in G^m;$$

$$u = (u_1,\ldots,u_n), v = (v_1,\ldots,v_n), w = (w_1,\ldots,w_n) \in G^n;$$

$$\rho_1,\rho_2,\rho_3 \in S_m; \sigma_1,\sigma_2,\sigma_3 \in S_n; \text{ and}$$

$$\theta_1,\theta_2,\theta_3 \in \mathrm{Aut}(G^0).$$

Then

$$\begin{aligned}
&((g,u,\rho_1,\sigma_1,\theta_1) * (h,v,\rho_2,\sigma_2,\theta_2)) * (k,w,\rho_3,\sigma_3,\theta_3)\\
=&((g_{1\rho_2}\theta_2 h_1,\ldots,g_{m\rho_2}\theta_2 h_m),(u_{1\sigma_2}\theta_2 v_1,\ldots,u_{n\sigma_2}\theta_2 v_n),\rho_2\rho_1,\sigma_2\sigma_1,\theta_1\theta_2) * (k,w,\rho_3,\sigma_3,\theta_3)\\
=&((g_{1\rho_3\rho_2}\theta_2\theta_3 h_{1\rho_3}\theta_3 k_1,\ldots),(u_{1\sigma_3\sigma_2}\theta_2\theta_3 v_{1\sigma_3}\theta_3 w_1,\ldots),\rho_3\rho_2\rho_1,\sigma_3\sigma_2\sigma_1,\theta_1\theta_2\theta_3)\\
=&(g,u,\rho_1,\sigma_1,\theta_1) * ((h_{1\rho_3}\theta_3 k_1,\ldots,h_{m\rho_3}\theta_3 k_m),(v_{1\sigma_3}\theta_3 w_1,\ldots,v_{n\sigma_3}\theta_3 w_n),\rho_3\rho_2,\sigma_3\sigma_2,\theta_2\theta_3)\\
=&(g,u,\rho_1,\sigma_1,\theta_1) * ((h,v,\rho_2,\sigma_2,\theta_2) * (k,w,\rho_3,\sigma_3,\theta_3))
\end{aligned}$$

and so this multiplication is associative. Thus we have a group. Next we show that the multiplication is compatible with the proposed action:

$$\begin{aligned}
&(p_{i,j}) \cdot ((g,u,\rho_1,\sigma_1,\theta_1) * (h,v,\rho_2,\sigma_2,\theta_2))\\
=&(p_{i,j}) \cdot (g_{1\rho_2}\theta_2 h_1,\ldots,g_{m\rho_2}\theta_2 h_m),(u_{1\sigma_2}\theta_2 v_1,\ldots,u_{n\sigma_2}\theta_2 v_n),\rho_2\rho_1,\sigma_2\sigma_1,\theta_1\theta_2)\\
=&(h_i^{-1} g_{i\rho_2}^{-1}\theta_2 p_{i\rho_2\rho_1,j\sigma_2\sigma_1}\theta_1\theta_2 u_{j\sigma_2}\theta_2 v_j)\\
=&(g_i^{-1} p_{i\rho_1,j\sigma_1}\theta_1 u_j) \cdot (h,v,\rho_2,\sigma_2,\theta_2)\\
=&((p_{i,j}) \cdot (g,u,\rho_1,\sigma_1,\theta_1)) \cdot (h,v,\rho_2,\sigma_2,\theta_2).
\end{aligned}$$

It is clear that the identity fixes any matrix in $M_{m\times n}(G^0)$, therefore this is a group action. □

Herein we will neglect to redefine the homomorphism ψ or to use the subscript $\rtimes_\psi$ when referring to semidirect products of the form $(G^m \times G^n) \rtimes_\psi (S_m \times S_n \times \mathrm{Aut}(G^0))$. We will simply write $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ and the reader should assume that the semidirect product is defined using the ψ from Equation 2.3.

When it comes to implementing a method for counting orbits of this action knowing the kernel will allow us to save time. Before we conclude this section, we will describe the kernel of this action.

Lemma 2.2.2. *Let G be a group, and let $m,n>0$ be integers. Then the kernel of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ with respect to the action defined in Equation 2.1 is*

$$\{((g,\ldots,g),(g,\ldots,g),1_{S_m},1_{S_n},x\mapsto gxg^{-1}) : g\in G\}.$$

Proof. It is clear that for all $g\in G$ the element

$$((g,\ldots,g),(g,\ldots,g),1_{S_m},1_{S_n},x\mapsto gxg^{-1})$$

is in the kernel. Therefore we must show that all elements of the kernel are of this form.

Let $x=((g_1,\ldots,g_m),(u_1,\ldots,u_n),\rho,\sigma,\theta)$ be an element of the kernel, i.e. it fixes all matrices in $M_{m\times n}(G^0)$. In particular, the matrix where every entry is equal to the identity of G is fixed by x. If this matrix is fixed by x then $g_i^{-1}1\theta u_j = 1$ for all $i\in\mathbf{m}$ and $j\in\mathbf{n}$. It follows that $1\theta = g_i u_j^{-1}$ and so $g_i = u_j$ for all i,j. Therefore x is of the form $x=((g,\ldots,g),(g,\ldots,g),\rho,\sigma,\theta)$ for some $g\in G$. Next, consider the matrix where every element is equal to some element h of G. If this matrix is fixed by x then $g^{-1}h\theta g = 1$ must hold. It follows that $h\theta = ghg^{-1}$ and so $g_i = u_j$. In fact this must be true for all $h\in G$ and so θ is the automorphism which conjugates by g^{-1}. Next, consider the matrix which has all entries equal to 1 except the first entry of the first row which is equal to 0. Then if x fixes this matrix it must be the case that ρ fixes 1 and σ fixes 1. By taking the generalisation of this example we see that ρ and σ must fix all rows and all columns if x is in the kernel. This completes the proof. □

We have now shown how to construct a group whose orbits correspond with the isomorphism classes of 0-simple semigroups of type (G,m,n). In the next section we will investigate alternative representations of matrices which will simplify the corresponding induced action of this class of groups.

2.3 Matrix representations

In this section we will introduce a novel representation for matrices over 0-groups which will allow us to enumerate orbits more effectively. This representation involves replacing the space of $m \times n$ matrices with the space of size m multisets of row vectors (which have length n). Roughly speaking, if we can represent the group action we defined in the previous section on a smaller space than $M_{m\times n}(G^0)$ then it will be computationally easier to determine the number of orbits. Moreover our representation will be such that the S_m and G^m factors of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ act trivially, further simplifying our computations.

We begin by defining a representation of matrices as functions. Let F be an $m \times n$ matrix with entries from the 0-group G^0. Then we will show how a function $f : (G^0)^n \to \mathbb{N}$ can represent F, up to permutations of the rows. An element x of $(G^0)^n$ is an n-tuple of entries from G^0 and we will consider it as a row of a matrix with n columns. For each x in $(G^0)^n$ the image $f(x) \geq 0$ should be interpreted as the number of times the row x appears in the matrix which f represents. The sum of the images $w(f)$ is called the weight of f:

$$w(f) = \sum_{x \in (G^0)^n} f(x)$$

and it corresponds with the total number of rows of the matrix F which f represents. Since this representation is independent of the ordering of the rows, the action of S_m which permutes the rows of matrices in $M_{m\times n}(G^0)$ becomes trivial in this representation. The collection of all functions representing matrices in $M_{m\times n}(G^0)$ in this way corresponds with S_m orbits of $M_{m\times n}(G^0)$ and is significantly smaller than the whole of $M_{m\times n}(G^0)$.

Example 2.3.1. Consider the 3×2 matrices over G^0 where $G = \{1, g\}$ is a cyclic group of order 2. The set $(G^0)^2$ has 9 elements and is equal to:

$$\{(0,0),(0,1),(0,g),(1,0),(1,1),(1,g),(g,0),(g,1),(g,g)\}.$$

Define the following three matrices

$$F_1 = \begin{pmatrix} 1 & 1 \\ 1 & g \\ g & 0 \end{pmatrix}, \; F_2 = \begin{pmatrix} 1 & g \\ g & 0 \\ 1 & 1 \end{pmatrix}, \; F_3 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Then both F_1 and F_2 are represented by the function from $(G^0)^2$ into $\mathbb{N}$ which maps $(1,1),(1,g)$, and $(g,0)$ to 1, and all other elements of $(G^0)^2$ to 0. The matrices F_1 and F_2 have the same rows but in a different order but this representation ignores the ordering of rows. The matrix F_3

is represented by the function from $(G^0)^2$ into $\mathbb{N}$ which maps $(1,1)$ to 3, and maps all other elements of $(G^0)^2$ to 0.

We can further simplify the action by continuing to modify this representation. The action of G on $(G^0)^n$ via

$$g \cdot (g_1, \ldots, g_n) = (g^{-1}g_1, \ldots, g^{-1}g_n),$$

for g in G and $(g_1, \ldots, g_n)$ in $(G^0)^n$, can be thought of as left inverse multiplication of a row of a matrix over G^0 with n columns. We will denote the orbits this action of G on $(G^0)^n$ by $(G^0)^n//G$. All isomorphism classes of (G,m,n) 0-simple semigroups can be represented by functions $(G^0)^n//G \to \mathbb{N}$ with weight m, since replacing the row of a matrix with another row of a matrix with a row from the same orbit of $(G^0)^n//G$ results in the matrix of an isomorphic Rees 0-matrix semigroup. We will denote the collection of all these functions representing $m \times n$ matrices over G^0 by

$$F_n^m(G^0) = \{f \in (G^0)^n//G \to \mathbb{N} : w(f) = m\}.$$

We note that each of these functions corresponds with an orbit of $M_{m\times n}(G^0)$ under the action of $G^m \rtimes S_m$ acting by row multiplication and permuting rows. Now, if we consider the induced action[6] of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ on the subsets of $M_{m\times n}(G^0)$ corresponding to $F_n^m(G^0)$, we see that the action of the G^m factor (which acts on the rows of matrices by left inverse multiplication) acts trivially, and the S_m factor also acts trivially. The representation of this induced action is a group isomorphic to $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$. Put another way, the action of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ on $F_n^m(G^0)$ is not faithful. The group $G^m \rtimes S_m$ is a normal subgroup and is a subset of the kernel of the induced action on $F_n^m(G^0)$. If we take the quotient with respect to $G^m \rtimes S_m$ we obtain a group isomorphic to $G^n \times (S_n \times \mathrm{Aut}(G^0))$. Thus we will find the isomorphism classes of 0-simple semigroups of type (G,m,n) by finding the orbits of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ with respect this action. Moreover, it should be computationally easier than working with $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$, a larger group, acting on $M_{m\times n}(G^0)$, a larger set.

Theorem 2.3.2. *The isomorphism classes of* 0*-simple semigroups of type* (G,m,n) *correspond to the orbits of*

$$G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$$

acting on $F_n^m(G^0)$ *via*

$$([(a_1, \ldots, a_n)])f^{((g_1,\ldots,g_n),\sigma,\theta)} = ([((a_{1\sigma}\theta)g_1, \ldots, (a_{n\sigma}\theta)g_n)])f$$

[6] See Section 1.8 for how we define an induced action.

Note that the product of two elements of this group is defined by

$$(\rho,(g_1,\ldots,g_n),\theta)(\sigma,(h_1,\ldots,h_n),\gamma)=(\sigma\rho,((g_{1\sigma})\gamma h_1,\ldots,(g_{n\sigma})\gamma h_n),\theta\gamma).$$

Proof. Clearly every $m\times n$ with entries from G^0 can be represented as an element of $F_n^m(G^0)$ and so every isomorphism class corresponds with at least one element of $F_n^m(G^0)$. On the other hand, each element of $F_n^m(G^0)$ represents a $G^m \rtimes S_m$ orbit of $M_{m\times n}(G^0)$ and thus an orbit of $F_n^m(G^0)$ contains all $G^m \rtimes S_m$ which have elements in the same orbits of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$. In other words, $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ orbits of $F_n^m(G^0)$ represent all matrices in the same $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbit of $M_{m\times n}(G^0)$. Therefore each orbit corresponds to a distinct isomorphism class of 0-simple semigroups.

We will now show that $G^n \rtimes (S_n \rtimes \mathrm{Aut}(G^0))$ acts on $F_n^m(G^0)$ in the manner described. Let $((g_1,\ldots,g_n),\rho,\theta),((h_1,\ldots,h_n),\sigma,\gamma)\in G^n \rtimes (S_n \times \mathrm{Aut}(G^0)$. Then if $f\in F_n^m(G^0)$ we have

$$\begin{aligned}
&([(a_1,\ldots,a_n)])(f^{((g_1,\ldots,g_n),\rho,\theta)})^{((h_1,\ldots,h_n),\sigma,\gamma)}\\
=&([((a_{1\rho})\theta g_1,\ldots,(a_{n\rho})\theta g_n)])f^{((h_1,\ldots,h_n),\sigma,\gamma)}\\
=&([(((a_{1\sigma\rho})\theta)\gamma(g_{1\sigma})\gamma h_1,\ldots,((a_{n\sigma\rho})\theta)\gamma(g_{n\sigma})\gamma h_n)])f\\
=&([(a_1,\ldots,a_n)])f^{(((g_{1\sigma})\gamma h_1,\ldots,(g_{n\sigma})\gamma h_n),\sigma\rho,\theta\gamma)}\\
=&([(a_1,\ldots,a_n)])f^{((g_1,\ldots,g_n),\rho,\theta)((h_1,\ldots,h_n),\sigma,\gamma)}.
\end{aligned}$$

It is clear that $(1_{G^n},1_{S_n},1_{\mathrm{Aut}(G^0)})$ fixes all functions in $F_n^m(G^0)$ too. Therefore this is a group action. The multiplication of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ is induced from the multiplication of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ given in Equation 2.2. □

We finish this section with an example involving the second representation of matrices as functions which we have introduced in this section.

Example 2.3.3. Again, let us consider the 3×2 matrices over G^0 where $G=\{1,g\}$ is a cyclic group of order 2. The set $(G^0)^2//G$ of orbits has 5 elements:

$$\{(0,0)\},\{(0,1),(0,g)\},\{(1,0),(g,0)\},\{(1,1),(g,g)\},\{(1,g),(g,1)\}.$$

Define the following two matrices

$$F_1=\begin{pmatrix}1&1\\1&g\\g&0\end{pmatrix},\ F_2=\begin{pmatrix}g&1\\g&g\\g&0\end{pmatrix}.$$

Then both F_1 and F_2 are represented by the function from $(G^0)^2//G$ into $\mathbb{N}$ which maps $\{(1,1),(g,g)\},\{(1,g),(g,1)\}$ and $\{(1,0),(g,0)\}$ to 1, and all other elements of $(G^0)^2//G$ to 0. The matrix F_2 can be produced by multiplying the first and second rows of F_1 by g and also swapping these rows. This representation ignores the ordering of rows and doesn't differentiate between rows in the same orbit in $(G^0)^2//G$.

For a different perspective, consider a function f from $(G^0)^2//G$ into $\mathbb{N}$ which maps $\{(1,1),(g,g)\}$ to 2, $\{(1,0),(g,0)\}$ to 1, and all other elements of $(G^0)^2//G$ to 0. Then the matrices in $M_{3\times 2}(G^0)$ which correspond with f are those with two rows from $\{(1,1),(g,g)\}$ and one row from $\{(1,0),(g,0)\}$. The rows can be in any order. There are 24 such matrices.

2.4 Pólya enumeration

In this section we introduce Pólya enumeration theory and show how to apply it to enumerate the orbits of the class of group actions introduced in the previous section. Choose a finite group G and integers $m,n>0$. In order to find the number of isomorphism classes of 0-simple semigroups of type (G,m,n) we will use the Pólya Enumeration Theorem to count the orbits of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ in its action on $F_n^m(G^0)$. We will present the Pólya Enumeration Theorem but first we must introduce the linked concepts of the cycle index of an element of a group, and the cycle index of a group itself.

Let G be a group of permutations of the set X. Then the *cycle index* of $g \in G$ is the following multivariate function:

$$z_g = t_1^{c_1(g)} \cdots t_n^{c_n(g)}$$

where n is the size of X and $c_i(g)$ denotes the number of i-cycles of g. The *cycle index* of the group G is the average of the cycle indices of its elements:

$$Z_G(t_1,\ldots,t_n) = \frac{1}{|G|}\sum_{g\in G} t_1^{c_1(g)} \cdots t_n^{c_n(g)}.$$

Recall that a weight function $w : Y \to \mathbb{N}$ assigns positive integer weights to the elements of some set Y. Furthermore, when we have a weight function w associated with a set Y and a function ϕ from some set X to Y we will say the weight $w(\phi)$ of the function ϕ is the sum over all elements x of X of the weights $w(x\phi)$ of their images. Now we present the (weighted) Pólya enumeration theorem (this theorem is covered in, for example, [5, Theorem 7.3] and [15, §2.4]).

Theorem 2.4.1. *Let X, Y be sets and let G be a group of permutations of X. Let Y^X denote the set of all functions from X to Y. Then the group G acts on Y^X via*

$$x(\phi g) = (xg)\phi$$

for all $x \in X$, $\phi \in Y^X$, and $g \in G$. Suppose that w is a weight function on Y and the function $f(t) = f_0 + f_1 t + f_2 t^2 + \cdots$ is defined such that f_i is the number of elements of Y of weight i. Then the number of orbits containing functions of weight t with respect to this action is given by

$$F(t) = Z_G(f(t), f(t^2), f(t^3), \ldots).$$

To use Theorem 2.4.1 to determine the number of isomorphism classes of 0-simple semigroups of type (G, m, n) we set $X = \mathbf{m}$ and $Y = (G^0)^n//G$. The permutation group which acts on X^Y is the representation of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ as a group of permutations of X^Y. We will set the weight $w(x)$ of x to equal x for all $x \in \mathbf{m}$. This will mean that the weight of $\phi \in X^Y$ is equal to the number of rows of the matrices in the class of matrices it represents. Furthermore, the subset of X^Y containing only functions having weight m is $F_n^m(G^0)$. Thus if we evaluate $F(t)$ we will find number of isomorphism classes of 0-simple semigroups of type (G, m, n).

However, in order to apply Theorem 2.4.1 we require the cycle index of the permutation representation of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$. Calculating the cycle index of a group effectively often uses the fact that two elements of the same conjugacy class have the same cycle index. Therefore a good method would involve:

(i) finding a set of representatives of the conjugacy classes of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$,

(ii) calculating the size of the conjugacy classes of these representatives, and

(iii) determining the cycle index of these representatives.

The following lemma gives us a method to find representatives of the conjugacy classes of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$. In particular, every conjugacy class will contain elements of a very particular form, which allows us to find representatives of every class by considering only a small subset of the complete set of elements of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$.

Lemma 2.4.2. *Let $x = (x_1, \ldots, x_n)$ be an element of $(G^0)^n$, let $\sigma \in S_n$ and let $\theta \in \mathrm{Aut}(G^0)$. Let $r_1, \ldots, r_d$ be representatives of the cycles of σ and let $l_1, \ldots, l_d$ be the lengths of the respective*

cycles. Write $y = (y_1, \ldots, y_n)$ *where* $y_i = 1$ *if* $i \notin \{r_1, \ldots, r_n\}$ *and otherwise*

$$y_{r_j} = \prod_{k=0}^{l_j - 1} x_{r_j \sigma^{-k}} \theta^{l-k}.$$

Then

$$[(\sigma, x, \theta)] = [(\sigma, y, \theta)]$$

and, in particular, every conjugacy class contains elements with at most d non-identity parts to x, where d is the number of cycles of σ*. Furthermore these d non-identity parts are all in distinct parts of x corresponding to the cycles of* σ*.*

Proof. Let $j \in \{1, \ldots, n\}$ then, if j is not fixed by σ, conjugating (σ, x, θ) by

$$y = (1, (y_1, \ldots, y_n), 1)$$

where $y_j = x_j^{-1}$ and $y_i = 1$ for any $i \neq j$ gives an element such that the jth component of the $(G^0)^n$ part is the identity of G and the $j\sigma$th component is $x_j \theta x_{j\sigma^{-1}}$. Now we repeat this process so that the result has more trivial entries in the G^n part. Assuming $j\sigma^2 \neq j$, we conjugate again but this time by $z = (1, (z_1, \ldots, z_n), 1)$ where $z_{j\sigma} = (x_{j\sigma^{-1}} \theta x_j)^{-1}$ and $z_i = 1$ for any $i \neq j$. The result is an element with both the jth and the $j\sigma$th entries of the G^n factor being trivial and the $j\sigma^2$th entry equal to $x^j \theta^2 x_{j\sigma^{-1}} \theta x_{j\sigma^{-2}}$. Applying this repeatedly we obtain an element where the G^n part is trivial in all entries of the cycle of σ containing j except for $j\sigma$ which is equal to

$$x_j \theta^{l-1} x_{j\sigma^{-1}} \theta^{l-2} \cdots x_{j\sigma^2} \theta x_{j\sigma}.$$

If we substitute j for $j\sigma^{-1}$ in this expression we obtain the form we expressed in the lemma's statement

$$\prod_{k=0}^{l_j - 1} x_{r_j \sigma^{-k}} \theta^{l-k} = x_{j\sigma^{-1}} \theta^{l-1} x_{j\sigma^{-2}} \theta^{l-2} \cdots x_{j\sigma} \theta x_j.$$

□

Although useful Lemma 2.4.2 is also lacking as it does not construct a unique representative for each conjugacy class. However it will allow us to better enumerate isomorphism classes, by reducing the number of candidates we need to consider to find unique representatives of conjugacy classes. We still need to determine the size of the conjugacy class of a representative. We also need a method for determining the cycle index of these representatives. In the sections that follow we will consider restrictions on G which can be tackled with better methods than the general case.

2.5 Enumerating regular matrices

In this short section we detail how to enumerate orbits of regular matrices. Thus far we have discussed how to enumerate the orbits of $M_{m\times n}(G^0)$ with respect to the action of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$. However these methods count both regular and non-regular matrices. Fortunately, we can deduce the number of these orbits which contain regular matrices[7] from the total number.

Lemma 2.5.1. *Let G be a group. Let $X(m,n)$ denote the number of orbits of $M_{m\times n}(G^0)$ with respect to the action of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$. Then the number of orbits containing regular matrices is equal to*

$$X(m,n) - X(m-1,n) - X(m,n-1) + X(m-1,n-1)$$

for $m,n > 1$.

Proof. An element of $M_{m\times n}(G^0)$ which is not regular can have:

(i) at least one row of zeros,

(ii) at least one column of zeros, or

(iii) satisfy both (i) and (ii).

Note that being regular, or satisfying (i), (ii), or (iii) are properties such that a matrix in an orbit with respect to the action of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ satisfies one of these properties if and only if all the other matrices in that orbit also satisfy that property. An orbit contains regular binary matrices exactly when it is not of type (i) and is not of type (ii). The total number of orbits of type (i) or (ii) is equal to the number which are of type (i) plus the number which are of type (ii) minus the number which are of type (iii). The number of type (iii) are subtracted because they are the intersection of the orbits of type (i) and type (ii). We will show the number of orbits of type (i), (ii), and (iii) are $X(m-1,n)$, $X(m,n-1)$, and $X(m-1,n-1)$ (respectively).

First, for any matrix which has a zero row, the orbit containing that matrix contains another matrix where the row of zeros is the last row. Without loss of generality, let $A = (a_{i,j}), B = (b_{i,j})$ be $m \times n$ binary matrices where the last row is all zeros. Let A', B' be created from A, B (respectively) by removing the last row, which is a row of zeros. Then A, B are in the same orbit $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ if and only if A', B' are in the same orbit of $(G^{m-1} \times G^n) \rtimes (S_{m-1} \times S_n \times \mathrm{Aut}(G^0))$.

[7] Note that an orbit contains a regular matrix if and only if all matrices in the orbits are regular.

To see this, let A, B be in the same orbit of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$. This is true if and only if there exists $(g, u, \rho, \sigma, \theta) \in (G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ such that $(a_{i,j}\theta) = (g_i^{-1} b_{i,j} u_j)$. Let $g' = (1g, \ldots, (m-1)g)$. If $m\rho = m$ then let $\rho' \in S_{m-1}$ be defined by $i\rho' = i\rho$ for all $1 \leq i \leq m-1$. Then $(g', u, \rho', \sigma, \theta)$ sends A' to B'. Otherwise $m\rho \neq m$ so the $m\rho$th row of A must be a row of zeros and the $m\rho^{-1}$th row of B must be a row of zeros. Let $\rho' \in S_{m-1}$ be defined by $(m\rho^{-1})\rho' = m\rho$ and $i\rho' = i\rho$ for $1 \leq i \leq m-1$ such that $i \neq m\rho^{-1}$. Then $(g', u, \rho', \sigma, \theta)$ sends A' to B'. Therefore A' and B' are in the same orbit of $(G^{m-1} \times G^n) \rtimes (S_{m-1} \times S_n \times \mathrm{Aut}(G^0))$.

For the reverse implication, let A', B' be in the same orbit of $(G^{m-1} \times G^n) \rtimes (S_{m-1} \times S_n \times \mathrm{Aut}(G^0))$. Then there exists $(g, u, \rho, \sigma, \theta) \in (G^{m-1} \times G^n) \rtimes (S_{m-1} \times S_n \times \mathrm{Aut}(G^0))$. which sends A' to B'. Let $g' = (1g', \ldots, (m-1)g', 1_G)$. Let $\rho' \in S_m$ such that $i\rho = i\rho'$ for $1 \leq i \leq m-1$ and $m\rho = m$. Then the action of $(\rho', \sigma) \in S_m \times S_n$ sends A to B. Thus A, B are in the same orbit of $S_m \times S_n$.

Thus there is a 1-1 correspondence between the orbits of $(G^{m-1} \times G^n) \rtimes (S_{m-1} \times S_n \times \mathrm{Aut}(G^0))$. in its action on $M_{m-1\times n}(G^0)$ and the orbits of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ acting on $M_{m\times n}(G^0)$ with at least one row of zeros. The number of the former is equal to $X(m-1, n)$ by definition. The number of orbits of $M_{m\times n}(G^0)$ of type (ii) can be shown to correspond to $X(m, n-1)$ with an analogous argument. The number of orbits of $M_{m\times n}(G^0)$ of type (iii) can be shown to correspond to $X(m-1, n-1)$ by combining the arguments used for enumerating the orbits of type (i) and (ii). □

Consequently, enumerating orbits of regular matrices with respect to the action of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ is possible whenever we can enumerate all the orbits. Although it requires us to enumerate the orbits of matrices with smaller dimensions, these cases should be significantly faster to compute as the difficulty scales with the matrix size.

2.6 Special cases

In this section we introduce and analyse how various special properties of the group of a 0-simple semigroup type allows for superior methods of enumeration.

2.6.1 The trivial group

Let **1** denote the trivial group. The isomorphism classes of type $(\mathbf{1}, m, n)$ 0-simple semigroups correspond to the equivalence classes of regular binary matrices where the equivalence is being equal up to row and column permutations. Specifically, if we let $\sim$ denote this equivalence then $P \sim Q$ if and only if $[p_{i,j}] = [q_{i\rho, j\sigma}]$ for some permutations $\rho \in S_m$ and $\sigma \in S_n$ of the rows and

columns, respectively. Therefore the number of $\sim$-classes of $m \times n$ binary matrices is equal to the number of orbits of $S_m \times S_n$ acting on the set of all binary matrices. These quantities can be determined [27] using the Polya Enumeration Theorem. The only requisite is the cycle index of the representation of the action of $S_m \times S_n$ on $m \times n$ binary matrices via permuting rows and columns. An expression for the cycle index of $S_m \times S_n$ in this action is known [26] and expressible in terms of the cycle indices of S_m and S_n. Using this expression to enumerate orbits then applying Lemma 2.5.1 to count the regular matrix orbits, the enumeration of this case is relatively easy for low order 0-simple semigroups.

When considering all isomorphism classes of 0-simple semigroups with order less than or equal to some integer k, it is seemingly the case $\mathscr{H}$-trivial[8] 0-simple semigroups account for the vast majority of isomorphism classes. For evidence see the tables in Section 2.7 and for each order compare the number of isomorphism classes of $\mathscr{H}$-trivial 0-simple semigroups to the total number of isomorphism classes of 0-simple semigroups. A finite zero 0-simple semigroup is $\mathscr{H}$-trivial if and only if it has type $(\mathbf{1}, m, n)$ where $m, n > 0$ can be any integers but the group must be trivial. Based on this observation, we make the following conjecture.

Conjecture 2.6.1. *Let k_i denote the number of isomorphism classes of 0-simple semigroups of order less than or equal to i and let t_i denote the size of the subset of these which are $\mathscr{H}$-trivial. Then the sequence*

$$\frac{t_i}{k_i} \to 1$$

as i tends to infinity.

Very roughly speaking, the author's intuition is that the number of orbits of $m \times n$ binary matrices is far greater than, say, the number of orbits of $m/|G| \times n$ matrices over G^0 for some group G of order dividing m. Furthermore the author believes the ratio of the former to the latter grows quickly with the parameters m, n, and $|G|$. Despite there potentially being many groups of a particular order, and many divisors of mn to choose from for group orders, the author believes that the sum of number of orbits from all these cases will still be relatively small compared to the binary matrix cases. As mentioned before, the tables in Section 2.7 support this belief.

Before we conclude this short section we note that there is an interesting sub-case. Congruence free semigroups, which correspond with binary matrices with all rows distinct and all columns distinct. This case is covered in depth in Chapter 4.

[8] A semigroup is said to be a $\mathscr{H}$-trivial if all of its $\mathscr{H}$-classes are of size one.

2.6.2 Groups with no outer automorphisms

An automorphism θ of a group G is said to be an *inner automorphism* if there is an element of g of G such that θ maps elements of G to their conjugate by g. The collection of all inner automorphisms forms the *inner automorphism group* of G which is denoted $\mathrm{Inn}(G)$. The *outer automorphism group* of G is the quotient of the automorphism group $\mathrm{Aut}(G)$ by the inner automorphism group $\mathrm{Inn}(G)$, and is denoted by $\mathrm{Out}(G)$. Examples of groups with no outer automorphism include but are not limited to: the symmetric groups (except S_6) and the automorphism groups of non-abelian simple groups. Other than S_2, these examples are all complete groups[9] as they also have trivial center.

For groups G where the outer automorphism group $\mathrm{Out}(G)$ is trivial we can apply some simplifications to our method of enumerating isomorphism classes of 0-simple semigroups of type (G,m,n). It turns out that if we consider the orbits of the subgroup

$$\{(g,\rho,1_{\mathrm{Aut}(G^0)}) : g \in G^n, \rho \in S_n\} \cong G^n \rtimes S_n$$

then these are the same as the orbits of the whole group $G^n \rtimes (S_n \times \mathrm{Aut}(G^0)$. We note that this subgroup is isomorphic to the wreath product $G \wr S_n$ and we will prefer writing $G \wr S_n$ to using semidirect product notation for the remainder of this section. The following lemma proves our assertion that orbits of $G \wr S_n$ correspond with isomorphism classes of 0-simple semigroups of type (G,m,n) when G has no outer automorphisms.

Lemma 2.6.2. *Let G be a group with trivial outer automorphism group. Then isomorphism classes of 0-simple semigroups of type (G,m,n) correspond to orbits of functions*

$$(G^0)^n//G \to \mathbb{N}$$

of weight m, with respect to the action of $G \wr S_n$ on $(G^0)^n//G$ via

$$([(x_1,\ldots,x_n)])f^{((g_1,\ldots,g_n),\rho)} = ([(x_{1\rho}g_1,\ldots,x_{n\rho}g_n)])f.$$

The kernel of this action is:

$$\{((g,g,\ldots,g),1_{S_n}) : g \in Z(G)\}.$$

Proof. This follows from Theorem 2.3.2 and the fact the inner automorphism $g \mapsto h^{-1}gh$ can be represented by the action of multiplying all rows by h^{-1} and multiplying all columns by h. We will show that two functions f_1, f_2 in $(G^0)^n//G \to \mathbb{N}$ are in the same orbit with

[9] A group is said to be *complete* if it has trivial outer automorphism group and trivial center.

respect to the action of $G \wr S_n$ if and only if they are in the same orbit with respect to the action of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0)$. The forward implication should be clear, if there is an element $((g_1,\ldots,g_n),\rho)$ of $G \wr S_n$ which sends f_1 to f_2 then the function $((g_1,\ldots,g_n),\rho,1_{\mathrm{Aut}(G^0)})$ sends f_1 to f_2 also.

For the reverse implication, let f_1, f_2 be functions in $(G^0)^n//G \to \mathbb{N}$ which are in the same orbit of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0)$. This holds if and only if there exists $((g_1,\ldots,g_n),\rho,\theta) \in G \rtimes (S_n \times \mathrm{Aut}(G^0)$ such that

$$([(x_1,\ldots,x_n)])f_1^{((g_1,\ldots,g_n),\rho,\theta)} = ([(x_1,\ldots,x_n)])f_2$$

for all $[(x_1,\ldots,x_n)] \in (G^0)^n//G$. Since $\mathrm{Out}(G)$ is trivial we deduce that there exists an element $h \in G$ such that θ maps elements of G to their conjugate by h. Therefore, for all $[(x_1,\ldots,x_n)] \in (G^0)^n//G$ we may show

$$\begin{aligned}([(x_1,\ldots,x_n)])g =& ([(x_1,\ldots,x_n)])f_1^{((g_1,\ldots,g_n),\rho,\theta)}\\ =& ([(h^{-1}x_{1\rho}hg_1,\ldots,h^{-1}x_{n\rho}hg_n)])f_1\\ =& ([(x_{1\rho}hg_1,\ldots,x_{n\rho}hg_n)])f_1\\ =& ([(x_1,\ldots,x_n)])f_1^{((hg_1,\ldots,hg_n),\rho)}\end{aligned}$$

using the fact that

$$[(x_1,\ldots,x_n)] = [(zx_1,\ldots,zx_n)]$$

for all $z \in G$. Thus f_1, f_2 are in the same orbit of $G \wr S_n$, as required.

The kernel can be deduced as a corollary of Lemma 2.2.2. □

The strategy in counting the case where the group has no outer automorphisms follows the same blueprint as the strategy for the general case which we presented in Section 2.4. We need conjugacy class representatives, the cycle index of each representative, and the size of each conjugacy class. However, the lack of the $\mathrm{Aut}(G^0)$ factor in the acting group leads to significant simplifications. This section culminates with Lemma 2.6.7 which is the author's method for determining the fix of an element of $G \wr S_n$ with respect to the action on $(G^0)^n//G$. We can use the fixes of elements of $G \wr S_n$ to determine the cycle indices of elements of $G \wr S_n$ with the method described in Proposition 2.6.4. The conjugacy classes of $G \wr S_n$ have been determined by James and Kerber [24, Section 4.2]. Conjugacy class representatives can be found using the method of Cannon and Holt [6]. The following lemma describes the aforementioned method of finding conjugacy class representatives as well as giving a formula for the size of the class they belong to.

Lemma 2.6.3. *Let G be a group and let S_n denote the symmetric group on n points. Let $g^{(1)},\dots,g^{(s)}$ denote representatives of the conjugacy classes of G. Let $x^{(1)},\dots,x^{(t)}$ denote the conjugacy class representatives of S_n. For some $1 \leq i \leq t$ denote by $c_1,\dots,c_k$ the cycles of $x^{(i)}$ such that $r_1,\dots,r_k$ are representatives of the respective cycles and $l_1,\dots l_k$ are the cycle lengths. Then the set A_i of all $(g,x^{(i)}) \in G \wr S_n$ such that*

(i) $jg \in \{g^{(1)},\dots,g^{(s)}\}$ when $j \in \{r_1,\dots,r_k\}$ and $jg = 1_G$ otherwise;

(ii) if $r_d g = g^{(a)}$, $r_e g = g^{(b)}$ and $l_d = l_e$ then $d \leq e$ implies $a \leq b$.

is such that $\cup_{i=1}^{t} A_i$ is a set of representatives of the conjugacy classes of $G \wr S_n$. The size of $[(g,x^{(i)})]$ is

$$|[(g,x^{(i)})]| = |g^{C_{S_n}(x^{(i)})}| \cdot |[x^{(i)}]| \cdot \prod_{j=1}^{k} (|G|^{l_j - 1} \cdot |[r_j g]|)$$

where $g^{C_{S_n}(x^{(i)})}$ is the orbit of g with respect to the centralizer of $x^{(i)}$ in S_n; $[x^{(i)}]$ is the conjugacy class of $x^{(i)}$ in S_n; and $[r_j g]$ is the conjugacy class of $r_j g$ in G.

Proof. Let us denote $\mathscr{A} = \cup_{i=1}^{t} A_i$. First we will show that every conjugacy class has a representative in the set $\mathscr{A}$. We think of the elements of $\mathscr{A}$ as canonical within their conjugacy class and will show how to canonicalise a generic element of $G \wr S_n$. The process of canonicalising will be an application of various conjugations until an element is transformed into canonical form.

Let $g = g(g_1,\dots,g_n)$ be an element of G^n and let $x \in S_n$. Then there exists $a \in \mathbf{t}$ such that x is conjugate to the conjugacy class representative $x^{(m)}$ of S_n. Let $y \in S_n$ be such that $x^{(m)} = y^{-1}xy$ and let $h = (h_1,\dots,h_n) \in G^n$ be defined so that

$$(h,x^{(m)}) = (1_{G^n}, y^{-1})(g,x)(1_{G^n}, y).$$

Note that $1_{G^n} = (1_G,\dots,1_G)$ and that (g,x) is conjugate to $(h,x^{(m)})$.

The element $(h,x^{(m)})$ of $G \wr S_n$ is best thought of in terms of the disjoint cycles of $x^{(m)}$ and the corresponding subsequences of h. When we mention the 'corresponding subsequence' we mean that the disjoint cycle $(\alpha_1 \alpha_2 \dots \alpha_z)$ corresponds with the subsequences containing $\alpha_1 h, \alpha_2 h, \dots$, and $\alpha_z h$. That is because, in the action which defines the multiplication of the wreath product, a permutation in S_n which contains the disjoint cycle $(\alpha_1 \alpha_2 \dots \alpha_z)$ acts on the corresponding subsequence $\alpha_1 h, \alpha_2 h, \dots$, and $\alpha_z h$ in the same way as the permutation which contains the disjoint cycle $(\alpha_1 \alpha_2 \dots \alpha_z)$ and fixes every other element of n. Consequently we can canonicalising h by canonicalising each subsequence in turn, as each one can be treated independently.

We begin canonicalising h by conjugating $(h,x^{(m)})$ until we obtain an element which satisfies condition (i). It will be convenient to define a map, $\gamma : G \to G$ which is any map such that for all f in G the element $f\gamma$ conjugates f to one of the representatives $\{g^{(1)},\dots,g^{(s)})\}$ of the conjugacy classes of G. As in the statement of the lemma, let $c_1,\dots,c_k$ be the disjoint cycles of $x^{(m)}$ such that $r_1,\dots,r_k$ are representatives of the respective cycles, and $l_1,\dots,l_k$ are their respective lengths. Let $i \in \mathbf{k}$. Assume $l_i = 1$. Then we conjugate $(h,x^{(m)})$ by $(h',1_{S_n})$ where $r_i h' = (r_i h)\gamma$, and $jh' = 1$ for $j \neq r_i$. Let $(h'',x^{(m)})$ denote the result of this conjugation. Then $r_i h'' \in \{g^{(1)},\dots,g^{(s)})\}$ as required.

On the other hand, assume $l_i > 1$. Let c_i be the cycle $(\alpha_1\alpha_2\dots\alpha_{l_i})$. Without loss of generality, assume $r_i = \alpha_1$. Conjugate $(h,x^{(m)})$ by $(h',1_{S_n})$ where

$$\alpha_z h' = \left(\prod_{j=z}^{l_i}(\alpha_j h)\right)^{-1},$$

and $jh' = 1$ for $j \notin \{\alpha_2,\dots,\alpha_{l_i}\}$. Let $(h'',x^{(m)})$ denote the result of this conjugation. Then $\alpha_j h'' = 1_G$ for $2 \leq j \leq l_i$ as required. However $r_i h'' = \prod_{j=1}^{l_i}(\alpha_j h)$ which is not necessarily one of the representatives $\{g^{(1)},\dots,g^{(s)})\}$ of conjugacy classes of G. Let $h''' \in G^n$ be such that $\alpha_j h''' = (r_i h'')\gamma$ for all $j \in \mathbf{k}$ and $jh''' = 1_G$ otherwise. Then conjugate $(h'',x^{(i)})$ by $(h''',1_{S_n})$. Let $(h'''',x^{(m)})$ denote the result of this conjugation. Then $r_i h'''' \in \{g^{(1)},\dots,g^{(s)})\}$ as required. Furthermore $\alpha_j h'' = 1_G$ for $2 \leq j \leq l_i$ still holds, as required.

Thus we make the cycle c_i satisfy condition (i). We repeat this process for all $i \in \mathbf{k}$ and obtain a conjugate of $(h,x^{(m)})$ which satisfies condition (i), which we will denote $(\kappa,x^{(m)})$.

Assume $(\kappa,x^{(m)})$ does not satisfy condition (ii). Then there exists $d,e \in \mathbf{k}$ such that $r_d\kappa = g^{(a)}$, $r_e\kappa = g^{(b)}$, $l_d = l_e$, $d < e$ and $a > b$. Let $(\alpha_1,\dots,\alpha_{l_d})$ denote the cycle c_d and let $(\beta_1,\dots,\beta_{l_d})$ denote the cycle c_e. Then the permutation $\zeta = (\alpha_1\beta_1)(\alpha_2\beta_2)\cdots(\alpha_{l_d}\beta_{l_d})$ commutes with $x^{(m)}$. Moreover if we conjugate $(\kappa,x^{(m)})$ by $(1_{G^n},\zeta)$ then, roughly speaking, the result has G^n factor where the subsequences corresponding to c_d and c_e have been swapped. Let $(\kappa',x^{(m)})$ denote the result of this conjugation. Then $r_d\kappa' = r_e\kappa = g^{(b)}$ and $r_e\kappa' = r_d\kappa = g^{(a)}$, so condition (ii) is now satisfied for r_d and r_e. We may repeat this process for any other cases where condition (ii) is not satisfied until we obtain a conjugate of $(\kappa,x^{(m)})$ which satisfies condition (ii). This concludes the proof that every element of $G \wr S_n$ is conjugate to an element of $\mathscr{A}$.

Let $(g,x^{(i)})$ be an element of $\mathscr{A}$. We will now determine the size of its conjugacy class. We have that conjugating by (h,y) in $G \wr S_n$ is equivalent to conjugating by $(h,1_{S_n})$ then by $(1_{G^n},y)$

since:

$$\begin{aligned}[(g,x^{(i)})] =& \{(h,y)^{-1}(g,x^{(i)})(h,y) : (h,y) \in G \wr S_n\} \\ =& \{(h^{-1}y^{-1},y^{-1})(g,x^{(i)})(h,y) : (h,y) \in G \wr S_n\} \\ =& \{(1_{G^n},y^{-1})(h^{-1},1_{S_n})(g,x^{(i)})(h,1_{S_n})(1_{G^n},y) : (h,y) \in G \wr S_n\}.\end{aligned}$$

For any (h,y) in $G \wr S_n$ we will denote

$$A_{(h,y)} = \{(\kappa^{-1},1^{G^n})(h,y)(\kappa,1_{G^n}) : \kappa \in G^n\}.$$

Then we have that every element of the conjugacy class of (h,y) can be expressed as an element of $A_{(h,y)}$ conjugated by an element of $G \wr S_n$ of the form $(1_{G^n},y)$. In terms of the conjugacy class of $(g,x^{(i)})$ this means:

$$[(g,x^{(i)})] = \{(1_{G^n},y^{-1})(h,x^{(i)})(1_{G^n},y) : (h,x^{(i)}) \in A_{(g,x^{(i)})} \text{ and } y \in S_n\}.$$

Note that $G \wr S_n$ is partitioned by the collection of all sets of the form $A_{(h,y)}$. Furthermore note that a group acts on itself by conjugation. From this action we can induce an action of the subgroup $1 \wr S_n = \{(1_{G^n},y) : y \in S_n\}$ of $G \wr S_n$ on the sets of the form $A_{(h,y)}$ via

$$(A_{(h,y)})^{(1,\rho)} = \{(1,\rho^{-1})(\kappa,\sigma)(1,\rho) : (\kappa,\sigma) \in A_{(h,y)}\}.$$

We will prove that the proposed action is closed in the sense that it will send a set $A_{(h,y)}$ to another set of this form. Let $(h,y) \in G \wr S_n$ and let $\rho \in S_n$. Then:

$$\begin{aligned}(A_{(h,y)})^{(1,\rho)} =& \{(1_{G^n},\rho^{-1})(\kappa,y)(1_{G^n},\rho) : (\kappa,y) \in A_{(h,y)}\} \\ =& \{(1_{G^n},\rho^{-1})(\kappa^{-1},1_{S_n})(h,y)(\kappa,1_{S_n})(1_{G^n},\rho) : \kappa \in G^n\} \\ =& \{(\kappa^{-1}\rho^{-1},\rho^{-1})(h,y)(\kappa,\rho) : \kappa \in G^n\} \\ =& \{(\kappa^{-1}\rho^{-1},1_{S_n})(1_{G^n},\rho^{-1})(h,y)(1_{G^n},\rho)(\kappa\rho^{-1},1_{S_n}) : \kappa \in G^n\} \\ =& \{(\kappa^{-1}\rho^{-1},1_{S_n})(h\rho^{-1},\rho^{-1}y\rho)(\kappa\rho^{-1},1_{S_n}) : \kappa \in G^n\} \\ =& \{(\kappa^{-1},1_{S_n})(h\rho^{-1},\rho^{-1}y\rho)(\kappa,1_{S_n}) : \kappa \in G^n\} \\ =& A_{h\rho^{-1},\rho^{-1}y\rho}\end{aligned}$$

It follows that conjugacy classes of $G \wr S_n$ are partitioned by sets of the form $A_{(h,y)}$. Thus the size of the conjugacy class $[(h,y)]$ in $G \wr S_n$ can be determined as the size of $A_{(h,y)}$ multiplied by the size of the orbit of $A_{(h,y)}$ with respect to the recently defined action of the subgroup $1 \wr S_n$ of

$G \wr S_n$. That is to say

$$|[(h,y)]| = |[A_{(h,y)}]| \cdot |A_{(h,y)}| \tag{2.4}$$

We will denote this orbit by $[A_{(h,y)}]$. The orbit-stabilizer theorem tells us the size of this orbit is $|S_n|$ divided by the size of the stabilizer of $A_{(h,y)}$. An element $(1,\rho) \in 1 \wr S_n$ is in the stabilizer of $A_{(h,y)}$ if and only if $h\rho^{-1} = h$ and $\rho^{-1}y\rho = y$. The latter is true precisely when ρ is in the centralizer $C_{S_n}(y)$ of y in S_n. Therefore the stabilizer of $A_{(h,y)}$ in $1 \wr S_n$ is a subgroup of $1 \wr C_{S_n}(h,y)$. Applying the orbit-stabilizer theorem again, we see that the size of this stabilizer is the size of $C_{S_n}(y)$ divided by the size of the orbit $h^{C_{S_n}(y)}$ of h under the action of $1 \wr C_{S_n}(y)$. Thus we have

$$\begin{aligned} |[A_{(h,y)}]| =& \frac{|S_n|}{|C_{S_n}(y)|/|h^{C_{S_n}(y)}|} \\ =& |[y]||h^{C_{S_n}(y)}| \end{aligned} \tag{2.5}$$

since $|S_n|/|C_{S_n}(y)|$ is equal to the size of the conjugacy class of y in S_n. To complete the proof we must show that

$$\left|[A_{(g,x^{(i)})}]\right| = \prod_{j=1}^{k} (|G|^{l_j-1} \cdot |[r_j g]|).$$

To see this we recall how we canonicalised the G^n factor of an element of $G \wr S_n$ in order to make it satisfy condition (i). Recall there are k cycles of $x^{(i)}$ called $c_1,\ldots,c_k$. Let us denote the cycle c_j by $(\alpha_{j,1},\ldots,\alpha_{j,l_j})$ where $\alpha_{j,1} = r_j$ is the representative we distinguished in the lemma statement. The process of canonicalising the G^n factor we described earlier first conjugates $(h,x^{(i)})$ to $(\kappa,x^{(i)})$, where $r_j\kappa = \prod_{z=1}^{l_j} \alpha_{j,z} h$, and $j\kappa = 1_G$ when $j \notin \{r_1,\ldots,r_k\}$. Next, $(\kappa,x^{(i)})$ is conjugated to the $(g,x^{(i)}) \in \mathscr{A}$ such that $r_j\kappa$ is in the G conjugacy class $[r_j g]$ for all $j \in \mathbf{k}$. Thus, the number of elements of $A_{(g,x^{(i)})}$ is the number of tuples $(h_1,\ldots,h_n) \in G^n$ satisfying the simultaneous equations

$$\{\prod_{z=1}^{l_j} h_{\alpha_{j,z}} \in [r_j g] : 1 \leq j \leq k\}. \tag{2.6}$$

The equation

$$h_{\alpha_{j,1}} \cdots h_{\alpha_{j,l_j}} \in [r_j g]$$

has $|G|^{l_j-1} \cdot |[r_j g]|$ solutions since for any values of $h_{\alpha_{j,1}},\ldots,h_{\alpha_{j,l_j-1}}$ the equation is satisfied if $h_{\alpha_{j,l_j}}$ is in the set

$$\{(h_{\alpha_{j,1}} \cdots h_{\alpha_{j,l_j-1}})^{-1}\mu : \mu \in [r_j g]\}.$$

Thus the number of solutions to the k equations in (2.6) is

$$\left|[A_{(g,x^{(i)})}]\right| = \prod_{j=1}^{k}(|G|^{l_j-1}\cdot|[r_jg]|) \tag{2.7}$$

as required. We substitute Equation 2.5 and Equation 2.7 into Equation 2.4 to obtain

$$|[(g,x^{(i)})]| = |[A_{(g,x^{(i)})}]|\cdot|A_{(g,x^{(i)})}| = (|g^{C_{S_n}(x^{(i)})}|\cdot|[x^{(i)}]|)\cdot\left(\prod_{j=1}^{k}(|G|^{l_j-1}\cdot|[r_jg]|)\right)$$

which completes the proof. □

The other ingredient we need to calculate the cycle index of $G\wr S_n$ with respect to the action on functions from $(G^0)^n//G$ to $\mathbb{N}$ is the cycle indices of the representatives of the conjugacy classes. We determine these via calculating the number of fixed points of each of these elements and the following proposition.

Proposition 2.6.4. *Let X be a set. Let $\mathrm{fix}(x)$ denote the number of points fixed by the permutation $x\in S_X$. Denote by z_i the number of i cycles of x. Firstly, $z_1=\mathrm{fix}(x)$ is clear. We can calculate the other z_a by knowing the number of fixed points of certain powers of x:*

$$a\cdot z_a = \mathrm{fix}(x^a) - \sum_{\{1\le d<a\,:\,d|a\}} dz_d$$

This allows us to calculate $\{k_d : d \text{ divides } |x|\}$ after calculating $\mathrm{fix}(d)$ for each divisor d of x and using the above formula. Then we will have deduced the cycle index:

$$z(g) = \prod_{d\,|\,|x|} s_d^{k_d}.$$

Proof. Every point fixed by x^a and will be in some d-cycle of x where d is a divisor of a and there will be $d\cdot k_d$ fixed points that were in d-cycles of x in each case. To determine the number of a-cycles we can subtract the number

$$\sum_{\{1\le d<a:d|a\}} dk_d$$

from the total number of fixed points of x^a then divide by a to count cycles instead of fixed points. □

Note that we can minimize the number of calculations required to determine the cycle indices of all conjugacy class representatives of $G\wr S_n$ by noticing that if x,y are elements of a

group such that $y = x^a$ then the values $\mathrm{fix}(x^{ab})$ and $\mathrm{fix}(y^b)$ are equal, with respect to any action. We can also save doing one calculation per conjugacy class since we know the size of the space our permutations act on. If we have a permutation x acting on a space X and know the number of fixed elements of powers of x for all but one of the divisors of $|x|$, then we can calculate the last quantity via

$$\sum_{d \text{ divides } |x|} \mathrm{fix}(x^d) = |X|.$$

Next we will show how to calculate $\mathrm{fix}(x)$ for an element x of $G \wr S_n$. First we describe the elements which are fixed.

Lemma 2.6.5. *Let $\sigma \in S_n$. Denote by $c_1, \ldots, c_k$ the cycles of σ such that $r_1, \ldots, r_k$ are representatives of the respective cycles and $l_1, \ldots l_k$ are the cycle lengths. Let $g = (g_1, \ldots, g_n) \in G^n$ be such that $g_i = 1_G$ if i is not in $\{r_1, \ldots, r_k\}$. Then $f = [(f_1, \ldots, f_n)] \in (G^0)^n//G$ is fixed by $x = (g, \sigma)$ if and only if there is an $h \in G$ such that;*

$$f_{r_i\sigma}^{-1} h^{l_i} f_{r_i\sigma} = g_{r_i} \text{ or } f_{r_i\sigma} = 0$$

for $1 \leq i \leq k$ and the other f_i ($i \notin \{r_1\sigma, \ldots, r_k\sigma\}$) are defined by:

$$\begin{array}{llll}
f_{r_1} = h^{-1} f_{r_1\sigma} g_{r_1}, & f_{r_1\sigma^{-1}} = h^{-1} f_{r_1} g_{r_1\sigma^{-1}}, & \ldots, & f_{r_1\sigma^{2-l_1}} = h^{-1} f_{r_1\sigma^{3-l_1}} g_{r_1\sigma^{2-l_1}} \\
f_{r_2} = h^{-1} f_{r_2\sigma} g_{r_2}, & f_{r_2\sigma^{-1}} = h^{-1} f_{r_2} g_{r_2\sigma^{-1}}, & \ldots, & f_{r_2\sigma^{2-l_2}} = h^{-1} f_{r_2\sigma^{3-l_2}} g_{r_2\sigma^{2-l_2}} \\
\vdots & \vdots & & \vdots \\
f_{r_k} = h^{-1} f_{r_k\sigma} g_{r_k}, & f_{r_k\sigma^{-1}} = h^{-1} f_{r_k} g_{r_k\sigma^{-1}}, & \ldots, & f_{r_k\sigma^{2-l_k}} = h^{-1} f_{r_k\sigma^{3-l_k}} g_{r_k\sigma^{2-l_k}}
\end{array}$$

In other words, by repeatedly applying $f_{i\sigma} g_i = h f_i$ starting with from the value $f_{r_i\sigma}$ in each cycle.

Proof. First, note that $f^x = f$ in $(G^0)^n//G$ if and only if there exists h in G such that $f^x = (h, h, \ldots, h) \cdot f$ in $(G^0)^n$. This is equivalent to the statement

$$\exists h \in G \text{ such that } f_{i\sigma} g_i = h f_i \text{ for all } 1 \leq i \leq n \tag{2.8}$$

Recall that $g_i = 1_G$ when $i \notin \{r_1, \ldots, r_k\}$. We now show for each $1 \leq i \leq n$ it follows from (2.8) that either $f_{r_i\sigma} = 0$ or:

$$\begin{aligned} g_{r_i} =& f_{r_i\sigma}^{-1} h f_{r_i} \\ =& f_{r_i\sigma}^{-1} h^2 f_{r_i\sigma^{-1}} \\ & \vdots \\ =& f_{r_i\sigma}^{-1} h^{l_i - 1} f_{r_i\sigma^{2-l_i}} \\ =& f_{r_i\sigma}^{-1} h^{l_i} f_{r_i\sigma^{1-l_i}} \\ =& f_{r_i\sigma}^{-1} h^{l_i} f_{r_i\sigma} \end{aligned}$$

Which completes the proof of the forward implication. For the reverse, we assume that we have an element $[(f_1, \ldots, f_n)]$ of $(G^0)^n//G$ such that either $f_{r_i\sigma}^{-1} h^{l_i} f_{r_i\sigma} = g_{r_i}$ or $f_{r_i\sigma} = 0$ for each $1 \leq i \leq k$, and $f_{i\sigma} g_i = h f_i$ for $i \notin \{r_1, \ldots, r_k\}$ is used repeatedly to define the other f_i from the values $\{f_{r_j\sigma} : j \in \mathbf{k}\}$ already defined. Recall that $g_i = 1$ if $i \notin \{r_1, \ldots, r_k\}$. Now we must prove that $f_{i\sigma} g_i = h f_i$ for all $1 \leq i \leq n$ however we only need to show $f_{r_i\sigma^2} g_{r_i\sigma} = h f_{r_i\sigma}$ for each $1 \leq i \leq k$ since the others hold by assumption. From our assumptions (when $f_{r_i\sigma} \neq 0$):

$$\begin{aligned} g_{r_i} =& f_{r_i\sigma}^{-1} h f_{r_i} \\ =& f_{r_i\sigma}^{-1} h^2 f_{r_i\sigma^{-1}} \\ & \vdots \\ =& f_{r_i\sigma}^{-1} h^{l_i - 1} f_{r_i\sigma^{2-l_i}} \end{aligned}$$

but we also know that $g_{r_i} = f_{r_i\sigma}^{-1} h^{l_i} f_{r_i\sigma}$ so this implies (note $f_{r_i\sigma^2} = f_{r_i\sigma^{2-l_i}}$):

$$f_{r_i\sigma}^{-1} h^{l_i} f_{r_i\sigma} = f_{r_i\sigma^2} = f_{r_i\sigma}^{-1} h^{l_i - 1} f_{r_i\sigma^{2-l_i}}.$$

Thus $f_{r_i\sigma^2} = f_{r_i\sigma^2} g_{r_i\sigma} = h f_{r_i\sigma}$ which completes the proof. □

It is important to note that we have really described the elements $f \in (G^0)^n$ such that $[f] \in (G^0)^n//G$ is fixed by some element of $G \wr S_n$. Consequently, care must be taken if we want to count the number of elements of $(G^0)^n//G$ in order not to count elements multiple times. Note that each element of $(G^0)^n//G$ is an orbit which contains $|G|$ elements of $(G^0)^n$, except $[(0, 0, \ldots, 0)]$ which has only one element. The following example demonstrates the process involved in Lemma 2.6.5 and the issue of over counting.

Example 2.6.6. Consider the element

$$(g,\sigma) = ((1\,2), 1_{S_3}, 1_{S_3}, (1\,2\,3), 1_{S_3}), (1\,2\,3)(4\,5))$$

in $S_3 \wr S_5$. We will construct the elements of $((S_3)^0)^n//G$ which are fixed by (g,σ). Lemma 2.6.5 tells us that $f = [(f_1, f_2, f_3, f_4, f_5)] \in ((S_3)^0)^n//G$ fixed by (g,σ) must satisfy

$$f_2^{-1} h^3 f_2 = (1\,2) \text{ or } f_2 = 0, \tag{2.9}$$

and

$$f_5^{-1} h^2 f_5 = (1\,2\,3) \text{ or } f_5 = 0 \tag{2.10}$$

for some element h of G. If $f_2 \neq 0$ then Equation 2.9 can only be satisfied if h^3 is conjugate to $(1\,2)$, in which case $h \in \{(1\,2), (1\,3), (2\,3)\}$. Note that if we were instead looking for h such that h^3 were conjugate to $(1\,2\,3)$ then there would be no solutions other than $f_2 = 0$. The possible solutions for Equation 2.9 are:

$$f_2 = 0$$

$$f_2 = 1_{S_3}, (1\,2) \text{ and } h = (1\,2);$$

$$f_2 = (2\,3), (1\,3\,2) \text{ and } h = (2\,3);$$

$$f_2 = (1\,3), (1\,2\,3) \text{ and } h = (1\,3).$$

If $f_5 \neq 0$ then Equation 2.9 can only be satisfied if h^2 is conjugate to $(1\,2\,3)$, in which case $h \in \{(1\,2\,3), (1\,3\,2)\}$. The possible solutions for Equation 2.10 are:

$$f_5 = 0$$

$$f_5 = 1_{S_3}, (1\,2\,3), (1\,3\,2) \text{ and } h = (1\,2\,3);$$

$$f_5 = (1\,2), (1\,3), (2\,3) \text{ and } h = (1\,3\,2).$$

Note that at least one of f_2, f_5 must be equal to 0 since there is no h which satisfies both h^3 is conjugate to $(1\,2)$ and h^2 is conjugate to $(1\,2\,3)$.

Once we choose a value for f_2 and f_5 we can repeatedly apply the formula $f_{i\sigma} g_i = h f_i$ (from Lemma 2.6.5) until we have determined the other values in f. A little consideration reveals that if $f_2 = 0$ then f_1, f_3 are also equal to 0. Similarly, if $f_5 = 0$ then f_4 is also equal to 0. Thus $[(0,0,0,0,0)]$ is fixed by (g,σ) and other fixed functions have the form $[(f_1, f_2, f_3, 0, 0)]$ or $[(0,0,0, f_4, f_5)]$. We will construct a couple of these.

Assume $f_5 = 0$. Let $h = (1\,2)$ and $f_2 = (1\,2)$. Then we obtain

$$f = [(1_{S_3}, 1_{S_3}, (1\,2), 0, 0)]$$

which is indeed fixed by (g, σ) since

$$\begin{aligned}[(1_{S_3}, 1_{S_3}, (1\,2), 0, 0)]^{(g,\sigma)} &= [((1\,2), (1\,2), 1_{S_3}, 0, 0)] \\ &= [((1\,2)(1\,2), (1\,2)(1\,2), (1\,2)1_{S_3}, (1\,2)0, (1\,2)0)] \\ &= [(1_{S_3}, 1_{S_3}, (1\,2), 0, 0)].\end{aligned}$$

For a second example let $f_5 = 0$, $h = (1\,3)$, and $f_2 = (1\,2\,3)$. Then we obtain

$$f = [((1\,2\,3), (1\,2\,3), (2\,3), 0, 0)]$$

which is indeed fixed by (g, σ) since

$$\begin{aligned}[((1\,2\,3), (1\,2\,3), (2\,3), 0, 0)]^{(g,\sigma)} &= [((1\,2\,3)(1\,2), (2\,3), (1\,2\,3), 0, 0)] \\ &= [((1\,3)(2\,3), (1\,3)(2\,3), (1\,3)(1\,2\,3), (1\,3)0, (1\,3)0)] \\ &= [((1\,2\,3), (1\,2\,3), (2\,3), 0, 0)].\end{aligned}$$

In fact, these examples are the same since

$$\begin{aligned}[((1\,2\,3), (1\,2\,3), (2\,3), 0, 0)] &= [((1\,3\,2)(1\,2\,3), (1\,3\,2)(1\,2\,3), (1\,3\,2)(2\,3), (1\,3\,2)0, (1\,3\,2)0)] \\ &= [(1_{S_3}, 1_{S_3}, (1\,2), 0, 0)].\end{aligned}$$

Furthermore, not just these two but all six of the solutions where $f_5 = 0$ and $f_2 \neq 0$, which have the form $[(f_1, f_2, f_3, 0, 0)]$, are equal in $(G^0)^n//G$. Similarly, all six of the solutions where $f_2 = 0$ and $f_5 \neq 0$, which have the form $[(0, 0, 0, f_4, f_5)]$, are equal in $(G^0)^n//G$. As mentioned earlier, $[(0, 0, 0, 0, 0)]$ is also fixed by (g, σ). The three elements of $(G^0)^n//G$ we have just mentioned are the only ones fixed by (g, σ).

Now we use Lemma 2.6.5 to determine a formula for the number of elements fixed by a conjugacy class representative of $G \wr S_n$.

Lemma 2.6.7. *Let $x = (g, \sigma)$ be an element of $G \wr S_n$ such that σ has cycles of length $l_1, \ldots, l_k$ with representatives $r_1, \ldots, r_k$ (respectively) and where $ig = 1_G$ whenever $i \notin \{r_1, \ldots, r_k\}$. Then*

the number of elements fixed by x with respect to the action on $(G^0)^n//G$ can be calculated by:

$$|G|\operatorname{fix}(x) = \sum_{\{[h]:h\in G\}} |[h]| \prod_{i=1}^{k} (|\{\alpha \in G : \alpha^{-1} h^{l_i} \alpha = g_{r_i}\}| + 1)$$

(note that either $|\{x \in G : x^{-1} h^{l_i} x = g_{r_i}\}|$ is equal to zero when $h^{l_i} \notin [g_{r_i}]$ otherwise it is equal to the size of the centralizer of g_{r_i} in G.)

Proof. Lemma 2.6.5 showed that elements fixed by x correspond to a choice of h and choices of $f_{r_1\sigma}, f_{r_2\sigma}, \ldots, f_{r_k\sigma} \in G$ (one for each cycle of σ) satisfying

$$f_{r_i\sigma}^{-1} h^{l_i} f_{r_i\sigma} = g_{r_i} \text{ or } f_{r_i\sigma} = 0$$

for $1 \leq i \leq k$. Unless $f_{r_i\sigma}$ is equal to 0 it is clear that h^{l_i} must be in the conjugacy class of g_{r_i}. In the case where $f_{r_i\sigma}$ is not 0, the number of choices for $f_{r_i\sigma}$ is equal to the size of the centralizer of h^{l_i} and depends only on the conjugacy class of h^{l_i}. We will count elements of $(G^0)^n$ such that their class in $(G^0)^n//G$ is fixed by x, which means we will count each class $|G|$ many times. To account for this we divide the total by $|G|$. We count the fix of x as follows:

$$\begin{aligned}
|G|\operatorname{fix}(x) =& |\{(h,(f_1,\ldots,f_k)) \in G \times (G^0)^k : f_{r_i\sigma}^{-1} h^{l_i} f_{r_i\sigma} = g_{r_i} \text{ or } f_{r_i\sigma} = 0 \text{ for each } i\}| \\
=& \sum_{h\in G} |\{(f_1,\ldots,f_k) \in (G^0)^k : f_{r_i\sigma}^{-1} h^{l_i} f_{r_i\sigma} = g_{r_i} \text{ or } f_{r_i\sigma} = 0 \text{ for each } i\}| \\
=& \sum_{h\in G} \prod_{i=1}^{k} |\{\alpha \in G^0 : \alpha^{-1} h^{l_i} \alpha = g_{r_i} \text{ or } \alpha^{-1} h^{l_i} \alpha = 0\}| \\
=& \sum_{\{[h]:h\in G\}} |[h]| \prod_{i=1}^{k} (|\{\alpha \in G : \alpha^{-1} h^{l_i} \alpha = g_{r_i}\}| + 1)
\end{aligned}$$

□

In order to use Lemma 2.6.7 for enumeration we need to be able to calculate two things: the conjugacy classes of G and for all $g \in G$ and $1 < l \leq n$ we need a way to determine $|\{h \in G : h^l = g\}|$, the number of lth roots of g. The number of mth roots of a permutation in the symmetric group is known [30, Theorem 1]. Otherwise, the literature [1, 3, 34] on roots of permutations appears to be concerned with number of permutations in the symmetric group which have mth roots rather than enumerating the mth roots of a given permutation.

Generally when counting the isomorphism classes of 0-simple semigroups of a given order, the tougher cases will be of type (G,m,n) where m and n are relatively large, rather than

the cases where G is large. Thus we will be able to get by with an unsophisticated method for enumerating roots without a significant impact on performance. We end this section by applying Lemma 2.6.7 to the case explored in Example 2.6.6.

Example 2.6.8. Recall that in Example 2.6.6 we determined that there were 13 elements of $((S_3)^0)^5//S_3$ fixed by the element

$$(g,\sigma) = ((1\,2), 1_{S_3}, 1_{S_3}, (1\,2\,3), 1_{S_3}), (1\,2\,3)(4\,5))$$

of $S_3 \wr S_5$. Lemma 2.6.7 tells us that

$$|S_3|\operatorname{fix}((g,\sigma)) = \sum_{\{[h]:h\in S_3\}} |[h]| \prod_{i=1}^{2} (|\{\alpha \in S_3 : \alpha^{-1} h^{l_i} \alpha = g_{r_i}\}| + 1)$$

Where $l_1 = 3$, $l_2 = 2$, $r_1 = 1$, and $r_2 = 2$. These variables take on their usual meanings for a conjugacy class representative of the form described in Lemma 2.6.3. That is to say, l_i denotes the length of the ith cycle of σ and r_i is a representative of the ith cycle of σ so that g_{r_i} is the entry of the tuple g which may be a non-identity element of S_3 - in this case $r_1 = 1$ and $r_2 = 2$ so that $g_{r_1} = (1\,2)$ and $g_{r_2} = (1\,2\,3)$. Note that the product ranges over $i \in \{1,2\}$ since σ is composed of two disjoint cycles.

We are summing over the conjugacy classes of S_3 which are $[1_{S_3}], [(1\,2)]$ and $[(1\,2\,3)]$. These classes have sizes $1, 3$ and 2, respectively. When $h = 1_{S_3}$ we have

$$\begin{aligned}
&|[1_{S_3}]|(|\{\alpha \in S_3 : \alpha^{-1} 1_{S_3} \alpha = (1\,2)\}| + 1)(|\{\alpha \in S_3 : \alpha^{-1} 1_{S_3} \alpha = (1\,2\,3)\}| + 1) \\
=&1(0+1)(0+1) \\
=&1.
\end{aligned}$$

Since $h^3 = h^2 = 1_{S_3}$ and 1_{S_3} is conjugate to neither $(1\,2)$ nor $(1\,2\,3)$. This corresponds to the solution $[(0,0,0,0,0)]$. When $h \in [(1,2)]$ we have

$$\begin{aligned}
&|[(1\,2)]|(|\{\alpha \in S_3 : \alpha^{-1} (1\,2) \alpha = (1\,2)\}| + 1)(|\{\alpha \in S_3 : \alpha^{-1} (1\,2) \alpha = (1\,2\,3)\}| + 1) \\
=&3(2+1)(0+1) \\
=&9.
\end{aligned}$$

Since $h^3 = h$ will be conjugate to $(1\,2)$ but not $(1\,2\,3)$, and the size of the centralizer in S_3 of an elements of $[(1,2)]$ is 2. This corresponds to the solutions of the form $[(a,b,c,0,0)]$ and also

counts the solution $[(0,0,0,0,0)]$ three more times. Finally, when $h \in [(1,2,3)]$ we have

$$\begin{aligned}&|[(1\,2\,3)]|(|\{\alpha \in S_3 : \alpha^{-1} 1_{S_3} \alpha = (1\,2)\}|+1)(|\{\alpha \in S_3 : \alpha^{-1}(1\,3\,2)\alpha = (1\,2\,3)\}|+1)\\ =&2(0+1)(3+1)\\ =&8.\end{aligned}$$

Since $h^3 = h$ will be conjugate to $(1\,2)$ but not $(1\,2\,3)$, and the size of the centralizer in S_3 of an elements of $[(1,2)]$ is 2. This corresponds to the solutions of the form $[(a,b,c,0,0)]$ and also counts the solution $[(0,0,0,0,0)]$ two more times. The result is

$$6\,\mathrm{fix}((g,\sigma)) = 1+9+8 = 18$$

so $\mathrm{fix}((g,\sigma)) = 3$. Note that the reason we counted $[(0,0,0,0,0)]$ a total of six times was deliberate since we knew we would count every other element of $((S_3)^0)^5//G$ a total of six times. This happens because each element (except $[(0,0,0,0,0)]$) of $((S_3)^0)^5//G$ is an orbit containing six element of $((S_3)^0)^5$ and we counted each of these six elements.

2.6.3 Abelian groups

The fundamental theorem of abelian groups tells us that an abelian group is a direct product of cyclic groups of prime power order. We envisage that if we first tackle the easier case of cyclic groups of prime power order then it may be possible to apply this case to solve the abelian case. The cyclic group of order p^k is isomorphic to the additive integers modulo p: $(\mathbb{Z}/p^k\mathbb{Z})^+$, and its automorphism group is isomorphic to the multiplicative integers modulo p: $(\mathbb{Z}/p^k\mathbb{Z})^\times$, where the automorphism corresponding to $j \in (\mathbb{Z}/p^k\mathbb{Z})^\times$ is $i \mapsto ij \mod p^k$. The following result is well known, for example [37, §5.2].

Lemma 2.6.9. *The group of multiplicative integers modulo n:*

$$(\mathbb{Z}/p^k\mathbb{Z})^\times$$

is cyclic when n is 1, 2, 4, p^k *or* $2p^k$ *for any odd prime p. Furthermore, this group is isomorphic to the cyclic group of order* $p^k - p^{k-1}$.

Herein we will write C_{p^k} to denote $(\mathbb{Z}/p^k\mathbb{Z})^+$ and $Aut(C_{p^k})$ to denote $(\mathbb{Z}/p^k\mathbb{Z})^\times$. To avoid confusion between the identity element $0 \in (\mathbb{Z}/p^k\mathbb{Z})^+$ and the zero element of the semigroup $(C_{p^k})^0$ we will denote the latter by $\mathbf{0}$ for the remainder of this section. Unfortunately the method of determining the cycle index by determining cycle indices for representatives of conjugacy

classes which we have previously employed is of relatively little help here as conjugacy classes in an abelian group haves size one. Thus the conjugacy classes of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ would be relatively small. However, the property of elements having the same cycle index is more general than the property of being conjugate[10]. Therefore if we could easily determine classes of elements of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ which have the same cycle index we would be able to apply determine the cycle index of the whole group using the cycle indices of representatives of these classes together with the sizes of the classes.

This is something we have been unable to characterise. However we have been able to determine the size of the fix of an element of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$. This would be a key result, akin to Lemmas 2.6.5 and 2.6.7 in Section 2.6.2, if we could also determine classes of elements with equal cycle index.

Lemma 2.6.10. *Let $G = C_{p^k}$ and let $f = (f_1, \ldots, f_n)$ be an element of $(G^0)^n$. Let $\sigma \in S_n$. Let z denote the number of cycles of σ. Let $l_1, \ldots, l_z$ denote the lengths of the cycles of σ and choose representatives $r_1, \ldots, r_z$ from these cycles, respectively. Let $g = (g_1, \ldots, g_n)$ be an element of G^n satisfying $g_i = 1$ for any $i \notin \{r_1, \ldots r_z\}$. Let α be an element of $\mathrm{Aut}(G^0)$. Let $x = (\sigma, g, \alpha)$. Then the following hold.*

(i) *The element x fixes $[f]$ if and only if there exists $h \in G$ such that for all $1 \leq i \leq z$ either*

$$f_{r_i} = \mathbf{0} \text{ or } (1-\alpha^{l_i})f_{r_i} = g_{r_i} - (1+\alpha+\cdots+\alpha^{l_i-1})h$$

and the remaining f_i are defined by $f_i = h + f_{i\sigma^{-1}} - g_i$ for $i \notin \{r_1\sigma^{-1}, \ldots, r_z\sigma^{-1}\}$.

(ii) *The number of elements fixed by x can be determined using:*

$$\mathrm{fix}(x) = \frac{\sum_{h\in G}\prod_{i\in I_h}(1+\gcd((1-\alpha^{l_i}), p^k))}{|G|}$$

where I_h is defined to be

$$I_h = \{i \in \{1, \ldots, z\} : g_{r_i} - (1+\alpha+\cdots+\alpha^{l_i-1})h \in \langle(1-\alpha^{l_i})\rangle\}$$

and $\langle(1-\alpha^{l_i})\rangle$ denotes the subgroup of G generated by $(1-\alpha^{l_i})$.

[10] For example, consider the group generated by (12) and (34) which has a natural action on $\{1,2,3,4\}$. Then the elements (12) and (34) have the same cycle index with respect to this action yet this group is abelian so they cannot be conjugate.

Proof. We begin by proving (i). If an element $[(f_1,\dots,f_n)]$ is fixed by x then there must exist an $h \in G$ such that:

$$g_i + f_{i\sigma} = h + f_i\alpha \qquad \text{for } i \in \{1,\dots,n\}$$

Rephrasing and using the fact that $g_i = 1_G$ whenever $i \notin \{r_1,\dots,r_z\}$ we know that

$$\begin{aligned} g_{r_i} &= h + f_{r_i} - f_{r_i\sigma}\alpha && \text{for } i \in \{1,\dots,z\} \\ f_i &= f_{i\sigma}\alpha - h && \text{for } i \notin \{r_1,\dots,r_z\} \end{aligned}$$

holds. Then it follows that for each $1 \le i \le z$ either $f_{r_i\sigma^{-1}} = \mathbf{0}$ or

$$\begin{aligned} g_{r_i} &= h + f_{r_i} - f_{r_i\sigma}\alpha \\ &= h + f_{r_i} - (f_{r_i\sigma^2}\alpha - h)\alpha \\ &= h + h\alpha + f_{r_i} - f_{r_i\sigma^2}\alpha^2 \\ &= h + h\alpha + h\alpha^2 + f_{r_i} - f_{r_i\sigma^3}\alpha^3 \\ &\vdots \\ &= (h + h\alpha + \cdots + h\alpha^{l_i-1}) + f_{r_i} - f_{r_i\sigma^{l_i}}\alpha^{l_i} \\ &= (1 + \alpha + \cdots + \alpha^{l_i-1})h + f_{r_i} - f_{r_i}\alpha^{l_i} \\ &= (1 + \alpha + \cdots + \alpha^{l_i-1})h + (1 - \alpha^{l_i})f_{r_i} \end{aligned}$$

so that $(1-\alpha^{l_i})f_{r_i} = g_{r_i} - (1+\alpha+\cdots+\alpha^{l_i-1})h$. The latter can only happen when $g_{r_i} - (1 + \alpha + \cdots + \alpha^{l_i-1})h$ is in the subgroup $\langle(1-\alpha^{l_i})\rangle$ of G. This completes the proof of (i).

Now we prove (ii). Fix $h \in G$. It is clear that if $f_{r_i} \neq \mathbf{0}$ then $(1-\alpha^{l_i})f_{r_i} = g_{r_i} - (1+\alpha+\cdots+\alpha^{l_i-1})h$ can only happen if $g_{r_i} - (1+\alpha+\cdots+\alpha^{l_i-1})h$ is in the subgroup $\langle(1-\alpha^{l_i})\rangle$ of G. Given h and $(f_{r_1},\dots,f_{r_z})$ satisfying the above condition there is exactly one $(f_1,\dots,f_n)$ which agrees on the values of $(f_{r_1},\dots,f_{r_z})$ and which is fixed by x. Therefore we count the number of $f \in (G^0)^n$ such that $[f^x] = [f]$ by

$$\sum_{h\in G}\prod_{i\in I_h}(1+\gcd((1-\alpha^{l_i}),p^k))$$

although this counts the element $(\mathbf{0},\dots,\mathbf{0})$ $|G|$ many times. This number is equal to $|G|\operatorname{fix}(x)$ since there are $|G|$ many elements of $(G^0)^n$ in each element of $(G^0)^n//G$ except for $(\mathbf{0},\dots,\mathbf{0})$ which was also counted $|G|$ many times. □

2.6.4 Decomposable matrices

We term a $m \times n$ binary matrix *decomposable* if it is the adjacency matrix of a disconnected digraph. That is to say, we may partition $\mathbf{m}$ as $I_1, \ldots, I_\alpha$, and $\mathbf{n}$ as $J_1, \ldots, J_\alpha$ such that the (i, j)th entry is non-zero if and only if $(i, j) \in I_x \times J_x$ for some $x \in \alpha$. An example of such a matrix corresponding to a disconnected digraph with α connected components with $A_1, A_2, \ldots, A_\alpha$ representing non-decomposable sub-matrices and 0 representing all-zero sub-matrices is:

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 & 0 \\ 0 & A_2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & A_{\alpha-1} & 0 \\ 0 & 0 & \cdots & 0 & A_\alpha \end{pmatrix}$$

We will call these non-zero sub-matrices the *connected components* of a decomposable matrix. We will term a matrix with entries from a 0-group as *decomposable* if and only if the binary matrix obtained by replacing non-zero entries with 1 is decomposable. The following lemma displays how isomorphism between Rees 0-matrix semigroups constructed from decomposable matrices can be phrased in terms of the connected components.

Lemma 2.6.11. *Let G be a group. Let $m, n > 1$ be integers. Let P, Q be decomposable $m \times n$ matrices over the zero group G^0 with k connected components. Let $A_1, \ldots, A_k$ be the connected components of P, and let $B_1, \ldots, B_k$ be the connected components of Q. Let A_i have r_i rows and c_i columns. Then $S = \mathscr{M}^0[G; P]$ is isomorphic to $T = \mathscr{M}^0[G; Q]$ if and only if there exists a permutation $\psi \in S_k$ and an automorphism $\theta \in \mathrm{Aut}(G^0)$ such that for all $1 \leq i \leq k$ the matrices $A_i\theta$ and $B_{i\psi}$ are in the same orbit of the action of $(G^{r_i} \times G^{c_i}) \rtimes (S_{r_i} \times S_{c_i})$ on $r_i \times c_i$ matrices over the zero group G^0.*

Proof. Let $P = (p_{i,j})$ and $Q = (q_{i,j})$. By Theorem 1.6.2, the semigroups S and T are isomorphic if and only if there exists $(g, u, \rho, \sigma, \theta) \in (G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ such that $(p_{i,j}\theta) = (g_i^{-1} q_{i\rho, j\sigma} u_j^{-1})$. Let the row index set of A_i be denoted by R_i^A and let the column index set of A_i be denoted by C_i^A. Similarly, let R_i^B and C_i^B denote the row and column index sets of B_i. Then for all $i \in \mathbf{k}$ the sub-matrix of Q with indices in $R_i^A\rho \times C_i^A\sigma$ is in the same orbit of the action of $(G^{r_i} \times G^{c_i}) \rtimes (S_{r_i} \times S_{c_i})$ as $A_i\theta$. Since all other entries in P are 0 all other entries in Q are zero. Therefore the connected components of Q are exactly the images of connected components of P. □

As we saw in Section 2.6.2, there are definite advantages to reducing the problem from counting orbits of $(G^m \times G^n) \rtimes (S_m \times S_n)$ rather than $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$. This

result allows us to consider orbits of the former type for each connected component, and deduce the isomorphism type of the whole from these. Furthermore, the result takes a particularly nice form when we restrict our attention to Rees 0-matrix semigroups over a group with no outer automorphisms.

Corollary 2.6.12. *Let G be a group with no outer automorphisms. Let $m,n>1$ be integers. Let P,Q be decomposable $m\times n$ matrices over the zero group G^0 with k connected components. Let $A_1,\ldots,A_k$ be the connected components of P, and let $B_1,\ldots,B_k$ be the connected components of Q. For each $i\in\mathbf{k}$ let A_i have r_i rows and c_i columns. Then $\mathscr{M}^0[G;P]$ is isomorphic to $\mathscr{M}^0[G;Q]$ if and only if there exists a permutation $\psi\in S_k$ such that the Rees 0-matrix semigroups $\mathscr{M}^0[G;A_i]$ and $\mathscr{M}^0[G;B_{i\psi}]$ are isomorphic for all $i\in\mathbf{k}$.*

Proof. As we saw in Section 2.6.2, when G has no outer automorphisms the orbits of $(G^m\times G^n)\rtimes(S_m\times S_n)$ and $(G^m\times G^n)\rtimes(S_m\times S_n\times\mathrm{Aut}(G^0))$ in their actions on $m\times n$ matrices with entries from G^0 are the same. In this case, Lemma 2.6.11 tells us that A_i and $B_{i\psi}$ are in the same orbit of the action of $(G^{r_i}\times G^{c_i})\rtimes(S_{r_i}\times S_{c_i})$ on $r_i\times c_i$ matrices over the zero group G^0. Furthermore A_i and $B_{i\psi}$ will be in the same orbit of $(G^{r_i}\times G^{c_i})\rtimes(S_{r_i}\times S_{c_i})$ if and only if they form isomorphic Rees 0-matrix semigroups. □

These results can be used to reduce the number of calculations required to determine the number of isomorphism classes of 0-simple semigroups over decomposable matrices. As long as $m,n>1$ then for any group G there will be $m\times n$ regular matrices with entries from G^0 which are decomposable, and therefore isomorphism classes of 0-simple semigroups where this method is superior[11]. As a result, implementing a special method for 0-simple semigroups corresponding to Rees 0-matrix semigroups constructed to decomposable matrices could speed up enumeration of isomorphism classes of 0-simple semigroups of type (G,m,n) for all but a few trivial types. However, the time saved will be insignificant since the proportion of regular matrices of a given dimension which are also decomposable seems to be small. This theory would be of use to somebody who wants to enumerate isomorphism classes of 0-simple semigroups constructed from decomposable matrices based of a specific form. For example, say we wanted to enumerate up to isomorphism the 0-simple semigroups constructed from decomposable matrices of size 16×16 with four connected components $A_1,\ldots,A_4$ which are all 4×4 matrices. Then we only need to determine the $(G^4\times G^4)\rtimes(S_4\times S_4)$ orbits of 4×4 non-decomposable binary matrices and then apply Lemma 2.6.11. This is much easier than determining orbits of 16×16 matrices.

[11] For example, the matrix with 1_G on the diagonal and 0's elsewhere.

2.7 Results

In this section we give a brief overview of how the theory in Section 2.4 and Section 2.6.2 is applied to enumerate isomorphism classes of 0-simple semigroups. We also present tables of results produced from the implementation of these methods. The **GAP** code written by the author can be found in this GitHub repository https://github.com/ChristopherRussell/0-simple-semigroups.

2.7.1 Counting 0-simple semigroups

There is a good method (called **CycleIndex**) in GAP for finding the cycle index of group of permutations of $\mathbb{N}$. Seeing as we could not conceive a specialised method for finding the cycle index of elements of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ we will use **CycleIndex**, then apply Pólya Enumeration Theorem (Theorem 2.4.1) as described in Section 2.4. In order to apply this method we have to create the group in GAP which is not straightforward. To do this we need a representation of the action of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ on $(G^0)^n//G$ as a group of permutations on $\mathbb{N}$. Let ψ be a bijection from $(G^0)^n//G$ to $\{1,\ldots,|(G^0)^n//G|\}$. Then we can create an isomorphism $G^n \rtimes (S_n \times \mathrm{Aut}(G^0)) \to Sym(|(G^0)^n//G|)$ by

$$x \mapsto \psi^{-1} \circ x \circ \psi$$

for all $x \in G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$. Let $|G| = k$ and let γ be a bijection from G^0 to $\{1,\ldots,k+1\}$. Then our choice of ψ is the following function:

$$[(a_1,\ldots,a_n)] \mapsto (\sum_{i=1}^{n} |k+1|^{i-1}(a_i\gamma - 1)) + 1.$$

Next we find a set of generators for $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ and apply ψ to them. A possible choice for the set of generators for $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ is the union of the generators of the S_n and $\mathrm{Aut}(G^0)$ factors together with the generators of a single factor of G from G^n, i.e. $\{(g, 1_G, 1_G, \ldots, 1_G) : g \text{ is a generator of } G\}$. The S_n factor together with the generators of a single G factor of G^n generate the rest of the G^n factor. Once we have created a representation of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ in terms of $Sym(|(G^0)^n//G|)$ we apply **CycleIndex** then apply the Pólya Enumeration Theorem. Finally, we apply Lemma 2.5.1 to count the number of orbits which contain regular matrices, that is to say the orbits which correspond to isomorphism classes of 0-simple semigroups. The results of this process can be found in Tables 2.1-2.5.

In these tables of results we have provided the number of isomorphism classes of 0-simple semigroups of order n for all $1 \leq n \leq 130$. These calculations were performed in less than

six hours using a single process of **GAP** on a 2016 MacBook Pro with 2.6GHz quad-core Intel Core i7, 16GB of 2133MHz LPDDR3 RAM. These results could certainly be extended slightly further by running calculations for days on a similar machine, or by utilising parallel computing on multiple machines. The limiting factor as orders get larger is our reliance on the method **CycleIndex**. The cases that take the longest seem to be the ones where $M_{m\times n}(G^0)$ has the most orbits, with an exception when G is trivial. If we fix the order k by considering all (G,m,n) such that $k = |G| * |m| * |n| + 1$ then the number of orbits seems to be maximised when $|G| = 1$ and the difference between m and n is small[12]. The case where G is trivial is an exception because when G is trivial we can rely on a known expression for the cycle index of $S_m \times S_n$ as discussed in Section 2.6.1 and we do not need to use **CycleIndex**. As a result the cases which take longest for a specific order end up being those where the size of G is as small as possible but greater than one, and the difference between m and n is small. For example, the case $(C_2,8,8)$ took the longest of any case corresponding to 0-simple semigroups of order 129.

In the result tables we have also provided counts for three sub-classes of 0-simple semigroups. The count of groups with 0 is straightforward and is equal to the number of groups with order one less. The number of 0-simple semigroups which are simple with 0 counts those where the set of non-zero elements form a simple semigroup, and this corresponds to the number of simple semigroups of order one less[13]. Finally, the number of $\mathscr{H}$-trivial[14] 0-simple semigroups was included as evidence for Conjecture 2.6.1.

[12] In other words, when the matrices are close to being square. For example, the author would expect the $(G,2,8)$ case to have less orbits than the $(G,4,4)$ case

[13] Put another way, given any simple semigroup we can adjoin a zero element and obtain a 0-simple semigroup. If we remove the zero element from a 0-simple semigroup whose non-zero elements for a simple semigroup then we obtain a simple semigroup. If we adjoin a zero element then remove it we obtain the original semigroup, therefore simple semigroups of order k are in bijective correspondence with 0-simple semigroups of order $k+1$ whose non-zero elements form a simple semigroup.

[14] Recall that $\mathscr{H}$-trivial is equivalent to being of type (G,m,n) where G is trivial.

Order	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0-simple	0	1	3	3	9	3	18	3	33	23	30	3	151	3	46	189	294	3	487	3	1 397	664	90	3	5 648
$\mathscr{H}$-trivial	0	1	2	2	5	2	12	2	18	19	24	2	116	2	40	184	229	2	420	2	1 338	658	84	2	5 230
Simple with 0	0	1	3	3	7	3	10	3	17	7	10	3	29	3	10	9	46	3	32	3	31	10	10	3	102
Group with 0	0	1	1	1	2	1	2	1	5	2	2	1	5	1	2	1	14	1	5	1	5	2	2	1	15

Order	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
0-simple	3 841	118	1 917	11 716	3	45 398	3	31 996	4 503	186	204 185	285 790	3	226	9 706
$\mathscr{H}$-trivial	3 837	112	1 862	11 618	2	44 594	2	30 710	4 498	180	204 180	282 970	2	220	9 700
Simple with 0	7	10	22	32	3	56	3	166	9	10	9	178	3	10	10
Group with 0	2	2	5	4	1	4	1	51	1	2	1	14	1	2	2

Order	41	42	43	44	45	46	47	48	49	50	51	52
0-simple	1 116 191	3	3 583 160	3	399 960	3 881 357	318	3	30 271 239	29 610 810	15 214 393	35 305
$\mathscr{H}$-trivial	1 101 448	2	3 577 908	2	399 718	3 879 780	312	2	30 156 838	29 610 806	15 069 848	35 300
Simple with 0	153	3	78	3	36	83	10	3	661	7	80	9
Group with 0	14	1	6	1	4	2	2	1	52	2	5	1

Table 2.1 Number of isomorphism classes of various types of 0-simple semigroup of orders 1-52.

Order	53	54	55	56	57	58	59	60	61	62	63
0-simple	1 757 051	3	222 104 344	55 205 944	922 866 288	61 480	486	3	1 762 675 631	3	550
$\mathscr{H}$-trivial	1 756 696	2	222 070 638	55 205 938	922 263 078	61 474	480	2	1 757 017 984	2	544
Simple with 0	39	3	305	10	260	10	10	3	1 431	3	10
Group with 0	5	1	15	2	13	2	2	1	13	1	2

Order	64	65	66	67	68	69	70
0-simple	13 190 856 618	13 718 901 762	625 728 729	10 187 010 096	3	22 657 053	163 155
$\mathscr{H}$-trivial	13 190 836 800	13 715 458 490	625 728 724	10 186 888 516	2	22 656 374	163 150
Simple with 0	243	1 543	9	122	3	43	9
Group with 0	4	267	1	4	1	5	1

Order	71	72	73	74	75	76	77	78
0-simple	176 974 898 340	3	823 881 619 734	3	766	5 836 761 933	69 496 618	2 146 892 812 097
$\mathscr{H}$-trivial	176 878 810 270	2	823 708 035 742	2	760	5 834 602 820	69 495 720	2 146 892 812 092
Simple with 0	552	3	5 763	3	10	2 889	44	9
Group with 0	4	1	50	1	2	3	4	1

Table 2.2 Number of isomorphism classes of various types of 0-simple semigroup of orders 53-78.

Order	79	80	81	82	83	84	85
0-simple	356 407 493 734	3	19 580 066 324 253	20 677 459 330 076	930	3	26 419 833 132 046
$\mathscr{H}$-trivial	356 407 038 668	2	19 578 467 831 806	20 677 459 128 734	924	2	26 404 057 549 872
Simple with 0	162	3	14 566	736	10	3	14 646
Group with 0	6	1	52	15	2	1	15

Order	86	87	88	89	90	91	92
0-simple	46 093 054 629	1 018	552 177	464 292 341 808 739	3	2 094 085 011 265 979	260 383 999 114 719
$\mathscr{H}$-trivial	46 093 054 624	1 012	552 172	464 291 948 898 156	2	2 094 063 283 672 984	260 383 999 114 714
Simple with 0	9	10	9	923	3	67 314	9
Group with 0	1	2	1	12	1	10	1

Order	93	94	95	96	97	98	99
0-simple	514 299 114	789 124	1 206	315 901 054 341	10 256 563 628 288 892	3	2 597 046 185 315 768
$\mathscr{H}$-trivial	514 297 640	789 118	1 200	315 901 054 336	10 255 839 996 389 954	2	2 595 854 189 959 052
Simple with 0	48	10	10	9	185 084	3	8 219
Group with 0	4	2	2	1	231	1	5

Table 2.3 Number of isomorphism classes of various types of 0-simple semigroup of orders 79-99.

Order	100	101	102	103	104
0-simple	97 100 184 784 203 210	103 611 007 719 055 260	3	222 670 079 940 176	3
$\mathscr{H}$-trivial	97 100 184 782 573 478	103 610 729 300 331 610	2	222 670 075 515 736	2
Simple with 0	1 817	278 968	3	248	3
Group with 0	2	16	1	4	1

Order	105	106	107	108	109
0-simple	208 703 812 193 138 857	24 357 579 732 335 554	1 518	3	4 188 448 777 559 775 907
$\mathscr{H}$-trivial	208 703 805 552 811 252	24 357 568 048 361 550	1 512	2	4 188 419 057 295 337 524
Simple with 0	1 765	654 268	10	3	1 177 657
Group with 0	14	2	2	1	45

Order	110	111	112	113	114	115
0-simple	3	19 091 741 727 581 499 836	2 059 390	4 201 290 929 266 046 145	3	4 212 693 723 851 078
$\mathscr{H}$-trivial	2	19 091 738 417 871 913 466	2 059 384	4 200 967 305 301 010 852	2	4 212 693 711 897 014
Simple with 0	3	11 944	10	1 202 115	3	312
Group with 0	1	6	2	43	1	6

Table 2.4 Number of isomorphism classes of various types of 0-simple semigroup of orders 100-115.

Order	116	117	118	119	120
0-simple	10 412 895 280 125	6 565 200 607	167 692 379 464 165 700 555	1 866	1 814 914 300 200 409 639
$\mathscr{H}$-trivial	10 412 895 280 120	6 565 197 868	167 692 379 464 154 875 548	1 860	1 814 914 300 200 409 634
Simple with 0	9	55	4 507	10	9
Group with 0	1	5	4	2	1

Order	121	122	123	124	125
0-simple	1 700 378 443 600 053 588 142	1 745 061 194 503 344 181 720	1 990	3 619 077	13 963 742 950
$\mathscr{H}$-trivial	1 700 377 301 105 413 386 160	1 745 061 194 503 344 181 716	1 984	3 619 072	13 963 739 668
Simple with 0	28 032 758	7	10	9	56
Group with 0	47	2	2	1	4

Order	126	127	128	129	130
0-simple	51 505 577 788 824	6 275 790 700 072 159 443 717	3	1 203 342 802 955 800 797 357	4 708 924
$\mathscr{H}$-trivial	51 505 321 476 420	6 275 751 281 470 592 125 364	2	1 203 277 350 014 102 636 584	4 708 918
Simple with 0	2 759 049	65 212 924	3	10 195 549	10
Group with 0	5	16	1	2 328	2

Table 2.5 Number of isomorphism classes of various types of 0-simple semigroup of orders 116-130.

2.7.2 Counting 0-simple semigroups over a group with no outer automorphisms

We find representatives of conjugacy classes of $G^n \rtimes (S_n \times \mathrm{Aut}(G^0))$ and the sizes of their respective conjugacy classes using Lemma 2.6.3. We then apply Lemma 2.6.7 to determine the number of elements fixed by each conjugacy class representative. Proposition 2.6.4 allows us to determine the number of points fixed by each power of each conjugacy class representative. This gives us the cycle index of the representatives and hence we determine the cycle index of the whole group. The results of applying this method to certain (G,m,n) can be found in Tables 2.6-2.10. Note that our non-symmetric group examples of groups with no outer-automorphisms are of the form $C_p \rtimes \mathrm{Aut}(C_p)$ for prime p, and that $\mathrm{Aut}(C_p)$ is isomorphic to C_{p-1} in this case.

	$n=2$	3	4	5	6	7	8	9
$m=2$	5	11	25	46	86	145	243	384
3	11	92	729	6201	54167	452496	3506080	24967329
4	25	729	40 962	2 898 456	190 952 517	11 024 397 355	560 353 556 569	25 397 830 770 713

Table 2.6 Number of isomorphism classes of type (S_3, m, n) 0-simple semigroups.

	$n=2$	3	4	5	6	7	8	9
$m=2$	7	20	73	214	672	1 949	5 675	15 780
3	20	517	27 819	2 056 175	150 806 608	9 931 725 776	580 614 787 170	30 341 175 375 704

Table 2.7 Number of isomorphism classes of type ($C_5 \rtimes \mathrm{Aut}(C_5)$, m, n) 0-simple semigroups.

	$n=2$	3	4	5	6	7	8	9
$m=2$	7	23	91	321	1 209	4 394	15 792	54 378
3	23	886	96 845	12 017 173	1 305 588 619	122 429 025 261	10 068 441 798 523	737 217 168 687 444

Table 2.8 Number of isomorphism classes of type (S_4, m, n) 0-simple semigroups.

	$n=2$	3	4	5	6	7	8	9	10	11
$m=2$	7	24	119	681	4 564	29 980	185 116	1 054 956	5 558 100	27 208 678
3	24	2 966	1 216 493	458 291 381	144 917 687 503					

Table 2.9 Number of isomorphism classes of type ($C_7 \rtimes \mathrm{Aut}(C_7)$, m, n) 0-simple semigroups.

	$n=2$	3	4	5	6	7	8	9	10	11
$m=2$	9	53	649	10 436	192 351	3 356 683	53 343 781	766 623 761	10 025 157 483	120 223 352 757

Table 2.10 Number of isomorphism classes of type (S_5, m, n) 0-simple semigroups.

Chapter 3

Creating a database of 0-simple semigroups

Mathematicians have often tabulated sets of mathematical objects alongside their attempts to classify and count them. Before computers these lists were limited to what was practical to calculate by hand. In modern times databases have not only grown vastly in size. Computer databases linked to mathematical software are perhaps even more valuable than tables or lists as they allow users to instantiate and examine their entries with ease. A database provides an abundance of examples for use in research or teaching. For example, a researcher may test a hypothesis for a large number of cases in relatively little time.

In Chapter 2 we counted 0-simple semigroups up to isomorphism by counting the orbits of a particular group action of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ on the set of $m \times n$ matrices over a 0-group G^0. To create a database we need a representative of each isomorphism class, that is a Rees 0-matrix semigroup constructed from a representative of each of these orbits. Although finding a complete collection of representatives would allow one to count orbits, it is not the case that the method of counting orbits we used in the previous chapter yields representatives as a byproduct. The upshot of counting orbits as we did is that we could enumerate more larger cases than a method which also finds representatives could manage.

3.1 Finding a transversal

Given integers $m, n > 0$ and a finite group G our aim is to find exactly one representative of every orbit of $M_{m\times n}(G^0)$ with respect to the action of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$. We will call a set containing exactly one element of each orbit of an action a *transversal* of the orbits. Finding a representative of each orbit will be easier if there is a uniquely distinguished

element in every orbit. One way to pick unique representatives of every orbit would be to define a total order on matrices in $M_{m\times n}(G^0)$ and pick the minimal element of each orbit with respect to this order. To imagine a total order on the matrices of $M_{m\times n}(G^0)$ first consider interpreting the entries of a matrix as a sequence by reading each row from left to right, and reading the rows from top to bottom. Then these sequences could be ordered lexicographically, with respect to any total ordering of G^0, and this would be a total order on $M_{m\times n}(G^0)$. In general, if we have a method for finding unique representatives we will call these representatives *canonical*.

We will be using the **GAP** algorithm CANONICALIMAGE [25] to find canonical representatives of group orbits. This function takes an element of a set together with a group which acts upon that set and returns a canonical[1] element of the orbit of the input element. The crudest solution to use the algorithm CANONICALIMAGE to find unique representatives of every orbit of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ by applying this algorithm to every element of $M_{m\times n}(G^0)$ then discounting any repeated output. The time required to find a transversal in this manner will be closely related to the number of times CANONICALIMAGE is applied. The number of regular $m \times n$ matrices over a 0-group G^0 grows very quickly with the parameters $|G|$, m and n so applying CANONICALIMAGE to every element of $M_{m\times n}(G^0)$ quickly becomes infeasible. However we will still obtain a transversal by applying CANONICALIMAGE to a subset of $M_{m\times n}(G^0)$ as long as the subset contains at least one element from every orbit. For this reason, we will consider how to find small but easily determined subsets of $M_{m\times n}(G^0)$ which intersects every orbit. More precisely, we are looking for a subset for which the time saved by fewer applications of CANONICALIMAGE minus the time incurred to determine the subset is maximized.

The theory of *binary shapes* and *normalization* will allow us to find relatively small subspaces of $M_{m\times n}(G^0)$ which contain representatives of every orbit of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$. Moreover these subspaces are fairly easy to describe and iterate through. These topics will be the focus of Section 3.2 and Section 3.3, respectively.

We will also show how to find a transversal of 0-simple semigroups up to the equivalence being isomorphic or anti-isomorphic[2] since doing so is not drastically different from the isomorphism case. Let G be a finite group. Let P and Q be matrices over G^0. If either $\mathscr{M}^0[G;P]$ or $\mathscr{M}^0[G;P^T]$ is isomorphic to $\mathscr{M}^0[G;Q]$ then we will say $\mathscr{M}^0[G;P]$ is *equivalent up to isomorphism or anti-isomorphism* to $\mathscr{M}^0[G;Q]$. This equivalence is the join of the equivalences of isomorphism, and anti-isomorphism.

[1] Here canonical means that the output is the same for all elements of an orbit.

[2] An anti-isomorphism is a mapping between two semigroups which reverses the order of multiplication. That is to say, $\theta : S \to T$ is an anti-isomorphism if and only if $(st)\theta = (t)\theta(s)\theta$ for all $s,t \in S$. Two Rees 0-matrix semigroups may be anti-isomorphic but not isomorphic.

The motivation to find these semigroups up to isomorphism or anti-isomorphism would be that two anti-isomorphic semigroups behave almost identically. The key fact one needs to know about anti-isomorphisms between Rees 0-matrix semigroups is that the $\mathscr{M}^0[G;P]$ and $\mathscr{M}^0[G;P^T]$ are anti-isomorphic (where P^T denotes the transpose of P). Searching up to isomorphism or anti-isomorphism is only a non-trivial consideration in the case where the matrices are square, since we enumerate a traversal of the $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbits of $M_{m\times n}(G^0)$ for different G, m and n separately. When $m \neq n$ there are no distinct anti-isomorphic matrices in $M_{m\times n}(G^0)$. It is also worth noting that the orbits of $M_{n\times m}(G^0)$ are exactly the sets obtained by transposing the matrices in the orbits of $M_{m\times n}(G^0)$. Thus in the case of searching up to isomorphism we can determine the $n \times m$ case almost immediately from the $m \times n$ case.

Proposition 3.1.1. *Let G be a finite group. If S is a transversal of the $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbits of $M_{m\times n}(G^0)$ then $\{P^T : P \in S)\}$ is a transversal of the $(G^n \times G^m) \rtimes (S_n \times S_m \times \mathrm{Aut}(G^0))$ orbits of $M_{n\times m}(G^0)$.*

In the case where $m \neq n$ finding a transversal of the orbits of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ acting on $M_{m\times n}(G^0)$ will enumerate the 0-simple semigroups of type (G,m,n) and (G,n,m) up to isomorphism or anti-isomorphism. In the square case $M_{m\times m}(G^0)$ there may be anti-isomorphic matrices in different orbits of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$. In this case, to construct a group action which has orbits corresponding to equivalence up to isomorphism or anti-isomorphism of Rees 0-matrix semigroups we need to extend the action of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ by a group which acts on $M_{m\times m}(G^0)$ by transposing matrices. The extension will just be a group of order 2. However the operation of transposing a matrix does not commute with permuting rows, permuting columns, multiplying rows or multiplying columns. The extension we require is

$$((G^m \times G^m) \rtimes (S_m \times S_m \times \mathrm{Aut}(G^0))) \rtimes_\tau C_2 \tag{3.1}$$

where the homomorphism τ from C_2 into $\mathrm{Aut}((G^m \times G^m) \rtimes (S_m \times S_m \times \mathrm{Aut}(G^0)))$ sends the generator c of C_2 to the automorphism

$$(\{u_i\}_i, \{v_j\}_j, \phi, \chi, \theta) \mapsto (\{v_j\}_j, \{u_i\}_i, \chi, \phi, \theta). \tag{3.2}$$

which swaps $\{u_i\}_i \in G^m$ and $\{v_j\}_j \in G^m$ as well as swapping $\phi \in S_m$ and $\chi \in S_m$. The automorphism of θ of G^0 is unaffected.

3.2 Binary shapes

In this section we demonstrate how to find a relatively small subset of $M_{m\times n}(G^0)$ which contains a transversal. To do this we essentially identify a property which is satisfied by at least one element of every orbit. We then show how to generalise that method to one which has additional advantages in the case that G is a non-simple group, and this approach incorporates the theory of congruences on 0-simple semigroups.

Let G be a finite group and let P be in $M_{m\times n}(G^0)$. Let $\mathbf{1}$ denote the trivial group. We define the *binary shape* of P to be the $m\times n$ matrix $\mathscr{S}(P) = (b_{i,j})$ over $\mathbf{1}^0$ which has 0's in precisely the same locations as P and 1's elsewhere. That is to say:

$$b_{i,j} = \begin{cases} 0 & \text{if } p_{i,j} = 0, \\ 1 & \text{otherwise.} \end{cases}$$

We will say that two $m\times n$ matrices P,Q are *equivalent up to row and column permutations*, which we will denote by $P\equiv Q$, if there exist permutation $\phi \in S_m$ and $\chi \in S_n$ such that:

$$q_{i,j} = p_{i\phi,j\chi}$$

for all i in $\mathbf{m}$ and j in $\mathbf{n}$. The classes of the equivalence $\equiv$ are the orbits of the action of $S_m \times S_n$ which permutes the rows and columns of $m\times n$ matrices. It follows from Corollary 2.1.2 that two isomorphic Rees 0-matrix semigroup must have binary shapes which are equivalent up to row and column permutations. Denote by $M_{m\times n}(\mathbf{1}^0)$ the set of all $m\times n$ matrices over $\mathbf{1}^0$. Let us choose a transversal T of the $\equiv$-classes of $M_{m\times n}(\mathbf{1}^0)$. Then every $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbit of $M_{m\times n}(G^0)$ will contain an element which has binary shape equal to an element of T. It is unimportant how we choose a transversal of $M_{m\times n}(\mathbf{1}^0)$ however an example choice is as follows. We will use the total order on the two elements of $\mathbf{1}^0$ which has 0 being less than 1. Then we create a lexicographical order on $M_{m\times n}(\mathbf{1}^0)$ by deciding to read matrix entries from left to right, and by reading the row in turn from top to bottom. That is to say the matrix $(a_{i,j})$ is lexicographically less than $(b_{i,j})$ if there exists r,s such that $a_{r,s} < b_{r,s}$ but $a_{i,j} = b_{i,j}$ whenever $i < r$ or when $i = r$ and $j < s$. Each equivalence class of $M_{m\times n}(\mathbf{1}^0)$ will have a minimum element with respect to this ordering. For now, we shall call the least elements of the $\equiv$-classes of $M_{m\times n}(\mathbf{1}^0)$ with respect to the order we have just created the *lex-minimal binary shapes*. Then we have the following proposition:

Proposition 3.2.1. *Let G be a finite group and let $M_{m\times n}(G^0)$ be the space of all regular matrices over G^0. Then every isomorphism class of 0-simple semigroups contained in*

$$\{\mathscr{M}^0[G;P] : P \in M_{m\times n}(G^0)\}$$

contains an element whose binary shape is a lex-minimal binary shape of $M_{m\times n}(\mathbf{1}^0)$.

Proof. Let $S = (s_{i,j})$ be the binary shape of $P = (p_{i,j})$. Then there are row and column permutations ϕ and χ such that $(s_{i\phi,j\chi})$ is the lex-minimal binary shape of it's $\mathbf{1}^0$-class. Then $Q = (p_{i\phi,j\chi})$ has lex-minimal binary shape and $\mathscr{M}^0[G;Q]$ is isomorphic to $\mathscr{M}^0[G;P]$ by Corollary 2.1.2. □

A similar result is true for any transversal of the $\equiv$-classes of $M_{m\times n}(\mathbf{1}^0)$ but we will continue to consider the transversal containing the lex-minimal binary shapes for the sake of convenience. Let S be a lex-minimal binary shape of $M_{m\times n}(\mathbf{1}^0)$ and let $M_{m\times n}(G^0,S)$ denote the subset of $M_{m\times n}(G^0)$ consisting of the matrices with binary shape equal to S. Then every $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbit of $M_{m\times n}(G^0)$ has a non-empty intersection with the union

$$\bigcup_{S\in\mathbb{S}} M_{m\times n}(G^0,S) \tag{3.3}$$

where $\mathbb{S}$ is the set of all lex-minimal binary shapes from $M_{m\times n}(\mathbf{1}^0)$. This reduces our search space to one much smaller than the whole of $M_{m\times n}(G^0)$. It also allows us to treat $M_{m\times n}(G^0)$, which used to be a single case, as multiple smaller cases corresponding to the $\equiv$-classes of $M_{m\times n}(\mathbf{1}^0)$. It should be noted that this approach yields no improvement when G^0 is trivial because matrices from $M_{m\times n}(\mathbf{1}^0)$ are equal to their binary shape and two such matrices correspond to the same isomorphism class of Rees 0-matrix semigroups if and only if they are $\equiv$ related. We will see in Section 3.3 that we can further reduce the size of our search space in the case of matrices over a non-trivial group. A further advantage of this approach is that we can enumerate a transversal of equivalence classes of $M_{m\times n}(G^0,S)$ using a simpler group action, where P,Q are equivalent if and only if $\mathscr{M}^0[G;P]$ is isomorphic to $\mathscr{M}^0[G;Q]$. The acting group is smaller, the set acted upon is smaller, the orbits are smaller, and CANONICALIMAGE runs faster because it can more quickly determine the information it needs. This new group action is the action of the setwise stabilizer[3] of $M_{m\times n}(G^0,S)$ in $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$. The author's implementation only uses the theory described thus far.

Example 3.2.2. Let G be a finite group. We will consider binary shapes of 2×2 matrices over G^0 and what the sets $M_{m\times n}(G^0,S)$ might look like for certain binary shapes S. The following

[3]Let G be a group which acts on a set X, and let Y be a subset of X. Then the setwise stabilizer of Y is the subgroup of G containing all g such that Y^g is equal to Y.

sets are the 7 orbits of the 16 elements of $M_{2\times 2}(G^0)$ and the elements in the orbits are ordered lexicographically, with the lex-minimal element on the left.

$$\left\{\begin{pmatrix}1&1\\1&1\end{pmatrix}\right\},$$

$$\left\{\begin{pmatrix}1&1\\1&0\end{pmatrix},\begin{pmatrix}1&1\\0&1\end{pmatrix},\begin{pmatrix}1&0\\1&1\end{pmatrix},\begin{pmatrix}0&1\\1&1\end{pmatrix}\right\},$$

$$\left\{\begin{pmatrix}1&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&1\end{pmatrix}\right\},$$

$$\left\{\begin{pmatrix}1&0\\1&0\end{pmatrix},\begin{pmatrix}0&1\\0&1\end{pmatrix}\right\},$$

$$\left\{\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix}\right\},$$

$$\left\{\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right\},$$

$$\left\{\begin{pmatrix}0&0\\0&0\end{pmatrix}\right\}.$$

Now we will examine what the subsets of $M_{2\times 2}(G^0)$ which have a particular binary shape look like. Let us denote

$$S_1 = \begin{pmatrix}1&1\\1&1\end{pmatrix}, \text{ and } S_2 = \begin{pmatrix}1&1\\1&0\end{pmatrix}.$$

Then we have

$$M_{2\times 2}(G^0,S_1) = \left\{\begin{pmatrix}a&b\\c&d\end{pmatrix} : a,b,c,d \in G\right\},$$

and

$$M_{2\times 2}(G^0,S_2) = \left\{\begin{pmatrix}a&b\\c&0\end{pmatrix} : a,b,c \in G\right\}$$

Essentially, if S is a $m\times n$ binary matrix then $M_{m\times n}(G^0,S)$ contains all matrices where the 1's have been replaced by any element of G.

The theory of congruences on finite 0-simple semigroups and their connection to linked triples [19, §3.5] can be used to rephrase our ideas involving binary shapes. Moreover, we

will discover a generalisation of these ideas which can be applied when we are considering Rees 0-matrix semigroups over a non-simple group. Consider some Rees 0-matrix semigroup $\mathscr{M}^0[G;P]$ and let N be a normal subgroup of G. Then we may define a congruence ρ_N on $\mathscr{M}^0[G;P]$ by:

$$(i,a,\lambda)\rho_N(j,b,\mu) \iff i=j,\ \lambda=\mu \text{ and } ab^{-1}\in N. \tag{3.4}$$

This congruence corresponds to the linked triple $(N,\Delta_{\mathbf{m}},\Delta_{\mathbf{n}})$ where $\Delta_{\mathbf{m}}$ and $\Delta_{\mathbf{n}}$ are the diagonal relations on $\mathbf{m}$ and $\mathbf{n}$, respectively. The congruence ρ_N exists for all normal subgroups N of G. In particular, G is a normal subgroup of G and

$$(i,a,\lambda)\rho_G(j,b,\mu) \iff i=j \text{ and } \lambda=\mu \tag{3.5}$$

is the congruence on $\mathscr{M}^0[G;P]$ which corresponds to the linked triple (G,Δ_m,Δ_n). The congruence ρ_G relates to the earlier part of this section. Since the quotient of G by G is trivial we can think of ρ_G as mapping the Rees 0-matrix semigroup created from P to a Rees 0-matrix semigroup over $\mathbf{1}$ created from a matrix which is P with 1 replacing all elements of G and 0's unchanged. The latter matrix is essentially $S(P)$, the binary shape of P. The remainder of this chapter utilises the theory of linked triples to describe a superior approach for 0-simple semigroups over non-simple groups. For the unfamiliar we will give a brief overview of linked triples in relation to finite 0-simple semigroups. For a complete treatment see [19, §3.5]. Linked triples are a way of classifying the non-universal congruences of 0-simple semigroups[4]. Let G be a finite group and let $P=(p_{i,j})$ be a regular $m\times n$ matrix over G^0. Then a *linked triple* is a tuple composed of: a normal subgroup N of G, an equivalence $\mathscr{S}$ on $\mathbf{m}$, and an equivalence $\mathscr{T}$ on $\mathbf{n}$. Moreover, the equivalence $\mathscr{S}$ must be a subset of the equivalence

$$\varepsilon_{\mathbf{m}}=\{(i,j)\in\mathbf{m} : p_{i,k}=0 \iff p_{j,k}=0 \text{ for all } k\in\mathbf{n}\}, \tag{3.6}$$

and the equivalence $\mathscr{T}$ must be a subset of the equivalence

$$\varepsilon_{\mathbf{n}}=\{(i,j)\in\mathbf{n} : p_{k,i}=0 \iff p_{k,j}=0 \text{ for all } k\in\mathbf{m}\}. \tag{3.7}$$

There is a bijective correspondence [19, Theorem 3.5.8] between the non-universal congruences of $\mathscr{M}^0[G;P]$ and the set of linked triples given the parameters G, m, and n. There is also a formula to construct the congruence corresponding to a linked triple, and vice-versa, but we will not require the full theory.

[4] If we wanted to simultaneously cover the case of infinite semigroups then we would have to speak of completely 0-simple semigroups instead.

A linked triple of the form $(N, \Delta_{\mathbf{m}}, \Delta_{\mathbf{n}})$ where N is any normal subgroup of G, and where Δ_X denotes the trivial equivalence on X, corresponds with the congruence ρ_N of $\mathscr{M}^0[G;P]$ where:

$$(i,g,\lambda)\rho_N(j,h,\mu) \iff i=j,\ \lambda=\mu \text{ and } gh^{-1}\in N. \tag{3.8}$$

The author now defines a map which is helpful for constructing a semigroup isomorphic to the quotient of $\mathscr{M}^0[G;P]$ by ρ_N. Let

$$\pi_{G,N} : M_{m\times n}(G^0) \to M_{m\times n}((G/N)^0)$$

$$(p_{i,j}) \mapsto (p_{i,j}N)$$

where gN is a coset of N, i.e. an element of the quotient group G/N, and we define $0N$ to be equal to 0. We can consider ρ_N as a homomorphism from $\mathscr{M}^0[G;P]$ to $\mathscr{M}^0[G/N;\pi_{G,N}(P)]$ if we define

$$\begin{aligned}(i,g,\lambda)\rho_N &= (i,gN,\lambda)\\ 0\rho_N &= 0.\end{aligned} \tag{3.9}$$

The author has observed that every 0-simple semigroup constructed from a (regular) matrix in $M_{m\times n}(G^0)$ has $(N,\Delta_{\mathbf{m}},\Delta_{\mathbf{n}})$ as a linked triple[5]. As a consequence, a transversal of the $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbits of $M_{m\times n}(G^0)$ is contained in the union of $\pi_{G,N}$ preimages of a transversal of the $((G/N)^m \times (G/N)^n) \rtimes (S_m \times S_n \times \mathrm{Aut}((G/N)^0))$ orbits of $M_{m\times n}((G/N)^0)$. We now set out to prove the preceding statement regarding transversals.

We first consider a group homomorphism which has a special relationship with $\pi_{G,N}$. Let the map

$$\phi_{G,N} : (G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0)) \to ((G/N)^m \times (G/N)^n) \rtimes (S_m \times S_n \times \mathrm{Aut}((G/N)^0))$$

where if

$$x = ((g_1,\ldots,g_m),(u_1,\ldots,u_n),\rho,\sigma,\theta)$$

then

$$x\phi_{G,N} = ((g_1N,\ldots,g_mN),(u_1N,\ldots,u_nN),\rho,\sigma,\theta').$$

[5]Note that if P,Q are regular matrices in $M_{m\times n}(G^0)$ then it is not necessarily true that any linked triple of $S=\mathscr{M}^0[G;P]$ will be a linked triple of $T=\mathscr{M}^0[G;Q]$. For example let P be a 2×2 matrix with all entries equal to 1_G and let Q be a 2×2 matrix with a zero in the top left entry and 1_G's elsewhere. Then $(\{1_G\},\nabla_{\mathbf{m}},\nabla_{\mathbf{n}})$ is a linked triple of S but not of T. This is because in the context of T the equivalence $\nabla_{\mathbf{m}}$ is not a subset of the equivalence $\varepsilon_{\mathbf{m}}$ defined in Equation 3.6.

The automorphism θ' of $(G/N)^0$ is defined by $(gN)\theta' = (g\theta)N$. Our eventual aim is to show that $(P^x)\pi_{G,N} = (P\pi_{G,N})^{x\phi_{G,N}}$, which will allow us to link transversals of the two group actions. First we will demonstrate that θ' is well-defined, and that it is an automorphism. Let $a,b \in G$ be such that $aN = bN$. Then there exists an element g of N such that $a = bg$. It follows from the definition of θ' that

$$\begin{aligned} aN\theta' &= (bg)N\theta' \\ &= (bg)\theta N. \end{aligned}$$

Next, since θ is a homomorphism of G we have that

$$(bg)\theta N = ((b\theta)(g\theta))N.$$

Finally,

$$\begin{aligned} ((b\theta)(g\theta))N &= (b\theta N)(g\theta N) \\ &= (bN\theta')(gN\theta') \\ &= (bN\theta')(1N\theta') \\ &= (bN\theta')(1\theta N) \\ &= (bN\theta')(1N) \\ &= bN\theta'. \end{aligned}$$

Therefore θ' is well defined. The bijectivity of θ' follows immediately from the bijectivity of θ together with being well-defined. Lastly, for any $a,b \in G$ we have

$$\begin{aligned} (aNbN)\theta' &= (abN)\theta' \\ &= (ab)\theta N \\ &= ((a\theta)(b\theta))N \\ &= (a\theta N)(b\theta N) \\ &= (aN\theta')(bN\theta'). \end{aligned}$$

Therefore θ' is a homomorphism and thus an automorphism, as required. Now we prove the key relationship between $\pi_{G,N}$ and $\phi_{G,N}$.

Lemma 3.2.3. *Let G be a finite group and let N be a normal subgroup of G. Let $m,n>0$ be integers. Then for all $P \in M_{m\times n}(G^0)$ and $x \in (G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ we have:*

$$P^x \pi_{G,N} = (P\pi_{G,N})^{x\phi_{G,N}}.$$

Proof. Let $P=(p_{i,j})$ be a matrix from $M_{m\times n}(G^0)$. Let

$$x = ((g_1,\ldots,g_m),(u_1,\ldots,u_n),\rho,\sigma,\theta)$$

be an element of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$. Then

$$\begin{aligned}
(p_{i,j})^x \pi_{G,N} &= (g_i^{-1}(p_{i\rho,j\sigma}\theta)u_j)\pi_{G,N} \\
&= ((g_i^{-1}(p_{i\rho,j\sigma}\theta)u_j)N) \\
&= ((g_i^{-1}N)((p_{i\rho,j\sigma}\theta)N)(u_jN)) \\
&= ((g_i^{-1}N)((p_{i\rho,j\sigma}N)\theta')(u_jN)) \\
&= (p_{i,j}N)^{x\phi_{G,N}} \\
&= ((p_{i,j})\pi_{G,N})^{x\phi_{G,N}}
\end{aligned}$$

as required. □

We now use Lemma 3.2.3 to construct a relatively small subset of $M_{m\times n}(G^0)$ which contains a transversal of the $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbits.

Theorem 3.2.4. *Let G be a finite group and let N be a normal subgroup of G. Let $m,n>0$ be integers. Let T be a transversal of the $((G/N)^m \times (G/N)^n) \rtimes (S_m \times S_n \times \mathrm{Aut}((G/N)^0))$ orbits of $M_{m\times n}((G/N)^0)$. Then the union of preimages*

$$\bigcup_{P\in T} {\pi_{G,N}}^{-1}(P)$$

contains a transversal of the $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbits of $M_{m\times n}(G^0)$.

Proof. Let X be an orbit of $M_{m\times n}(G^0)$. We will show that an element of X is contained in

$$\bigcup_{P\in T} {\pi_{G,N}}^{-1}(P).$$

Let

$$Y = \{P\pi_{G,N} : P \in X\}$$

be the set of images of the matrices in X under $\pi_{G,N}$. Let Q be any element of X. Then we can show Y is an $M_{m\times n}((G/N)^0)$ orbit of $M_{m\times n}((G/N)^0)$ as follows:

$$\begin{aligned} Y &= \{P\pi_{G,N} : P \in X\} \\ &= \{(Q^x)\pi_{G,N} : x \in (G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))\} \\ &= \{(Q\pi_{G,N})^{x\phi_{G,N}} : x \in (G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))\} \\ &= \{(Q\pi_{G,N})^x : x \in ((G/N)^m \times (G/N)^n) \rtimes (S_m \times S_n \times \mathrm{Aut}((G/N)^0))\}, \end{aligned}$$

with the final step following from Lemma 3.2.3. Since Y is an orbit it has non-empty intersection with the transversal of orbits T. Thus if P is the matrix in the intersection of Y and T then the preimage $P{\pi_{G,N}}^{-1}$ has non-empty intersection with X. As X was arbitrary amongst orbits of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$, it follows that the union

$$\bigcup_{P\in T} {\pi_{G,N}}^{-1}(P)$$

contains at least one element from each $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbit of $M_{m\times n}(G^0)$. □

We can see how Proposition 3.2.1 follows from Theorem 3.2.4. First, if the normal subgroup N chosen is equal to the group G then $M_{m\times n}((G/N)^0)$ is equivalent[6] to $M_{m\times n}(\mathbf{1}^0)$. The map $\pi_{G,N}$ is the analogue of the map $\mathscr{S}$ which sends a matrix to its binary shape, the image of $\pi_{G,N}$ has the coset $1G$ where the image of $\mathscr{S}$ has 1. If we pick T to be the transversal of $M_{m\times n}((G/N)^0)$ which is formed of lex-minimal binary shapes then we obtain the result in Proposition 3.2.1.

3.3 Normalization

In this section we define normal matrices in the context of Rees 0-matrix semigroups. Our motivation comes from a normalization theorem which says that every finite 0-simple semigroup is isomorphic to some Rees 0-matrix semigroup constructed from a normal matrix. This result allows us to only consider normal matrices when finding a transversal of the isomorphism classes of 0-simple semigroups. Consequently we can reduce the number of times we call **CanonicalMatrix** and our computations will complete faster. For a rough idea of the improvement made, the set of $m \times n$ matrices with entries from a group G (with no zeros) is $|G|^{m+n-1}$

[6]Map a matrix over $\mathbf{1}^0$ to a matrix over $(G/G)^0$ be replacing the entries that are 1 with $1G$ and keeping entries that are 0 as 0.

times larger than the subset of these which are normal matrices[7]. The factor by which the subset of normal matrices of $M_{m\times n}(G^0)$ is smaller than the whole is harder to describe than the case with no zero entries. The factor is less than $|G|^{m+n-1}$ however it will remain significant.

Recall Theorem 1.6.2 and how left multiplying rows and right multiplying columns of a regular matrix with entries from a 0-group does not alter the isomorphism class of the corresponding Rees 0-matrix semigroup. In Chapter 2 we described the action of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ on $M_{m\times n}(G^0)$. We can also think of left multiplying rows and right multiplying columns in terms of the action of the subgroup $G^m \times G^n$ of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$. The theory of normal matrices is an attempt to find a normal form for the elements of each orbit of $G^m \times G^n$, though this form will generally not be unique. The process of finding a normal matrix in an orbit of $G^m \times G^n$ is what we mean when we say *normalization*.

The situation for matrices in $M_{m\times n}(G^0)$ with no 0 entries is straightforward and a good place to start. Consider $(p_{i,j}) \in M_{m\times n}(G^0)$ which has no 0 entries. Let $(q_{i,j})$ be the result of taking $(p_{i,j})$, left multiplying row i by $p_{i,1}^{-1}$ for all $i \in \mathbf{m}$, and right multiply column j by $p_{1,j}^{-1}$ for $1 < j \leq n$. The result satisfies $q_{1,j} = 1_G$ for all $j \in \mathbf{n}$ and $q_{i,1} = 1_G$ for all $i \in \mathbf{m}$. That is to say, $(q_{i,j})$ has the form:

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & q_{2,2} & q_{2,3} & \cdots & q_{2,n} \\ 1 & q_{3,2} & q_{3,3} & \cdots & q_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & q_{m,2} & q_{m,3} & \cdots & q_{m,n} \end{pmatrix}.$$

This demonstrates how any element of $M_{m\times n}(G^0)$ with no 0 entries is in the same $G^m \times G^n$ orbit as a matrix with at least $m+n-1$ entries equal to the identity of G. It is aesthetically pleasing for these identity entries to be all in the first row and first column but we can ensure $m+n-1$ identity entries in different locations too. We note that $m+n-1$ is the most identity entries we can guarantee for an element of the $G^m \times G^n$ orbit of an arbitrary $m \times n$ regular matrix with entries from an arbitrary 0-group. We demonstrate this with an example.

Example 3.3.1. Let $m,n > 0$ be integers and let $G = \{1, x, x^2, \ldots, x^{m^2n^2-1}\}$ be the cyclic group of order m^2n^2. Then consider the matrix $(p_{i,j})$ where $p_{i,j} = x^{ij}$:

$$\begin{pmatrix} x^1 & x^2 & \cdots & x^n \\ x^2 & x^4 & \cdots & x^{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x^m & x^{2m} & \cdots & x^{mn} \end{pmatrix}.$$

[7] In the case of matrices with no zeros, normal matrices are typically defined as matrices where all entries in the first row and first column are equal to the identity.

In order to reach a contradiction, assume there exists $(g_1,\dots,g_m)\in G^m$ and $(u_1,\dots,u_n)\in G^n$ such that $(q_{i,j})=(g_i^{-1}p_{i,j}u_j)$ has at least $m+n$ many entries which are equal to 1_G. Then there must be a,b,c,d where $1\leq a<b\leq m$ and $1\leq c<d\leq n$ and such that

$$q_{a,c}=q_{a,d}=q_{b,c}=q_{b,d}=1_G.$$

We will aim to show that $a=b$ or $c=d$ to reach a contradiction. We have:

$$g_a^{-1}x^{ac}u_c=1_G,$$

$$g_a^{-1}x^{ad}u_d=1_G,$$

$$g_b^{-1}x^{bc}u_c=1_G,$$

$$g_b^{-1}x^{bd}u_d=1_G.$$

Combining the former two equations, and combining the later two equations we obtain:

$$x^{ac}u_c=x^{ad}u_d,$$

$$x^{bc}u_c=x^{bd}u_d.$$

Multiplying the former by the inverse of the later we obtain:

$$x^{ac}x^{-bc}=x^{ad}x^{-bd},$$

which implies (since x has order m^2n^2)

$$abc^2=abd^2 \mod m^2n^2$$

since $a,b\in\mathbf{m}$ and $c,d\in\mathbf{n}$ we have that $1\leq abc^2\leq m^2n^2$ and $0\leq abd^2\leq m^2n^2$. Thus we can deduce that

$$abc^2=abd^2$$

which implies $c=d$. This is a contradiction, which proves there cannot have been $m+n$ many entries of $(q_{i,j})$ equal to 1_G.

The situation becomes more complicated when we allow for matrices with entries equal to 0. For instance, if we have a matrix $(p_{i,j})\in M_{m\times n}(G^0)$ with $p_{1,1}$ equal to 0 then there is no element its $G^m\times G^n$ orbit which has an identity in the first entry of the first row. Moreover, we are not guaranteed to have as many as $m+n-1$ identity entries. Most blatantly this can be

seen for a $m \times m$ square matrix where all entries not on the main diagonal are equal to 0:

$$\begin{pmatrix} x_1 & 0 & 0 & \cdots & 0 \\ 0 & x_2 & 0 & \cdots & 0 \\ 0 & 0 & x_3 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & x_m \end{pmatrix}.$$

This matrix is in the same $G^m \times G^m$ orbit as the matrix with all entries equal to 1_G on the main diagonal and equal to 0 elsewhere, and m identity entries is the most we can find in a matrix in this orbit. It will be important to the speed of our algorithms to work with matrices where we can guarantee as many identity entries as possible. We will define normal matrices differently than the definition found in [19, §3.4], henceforth we refer to this as the *standard definition* of normal matrices. We do this because the standard definition does not guarantee the maximum number of identity entries, nor does it lend itself well to analysis in this regard.

Example 3.3.2. We will demonstrate why the standard definition of normal matrices is inadequate for our purposes. The definition in [19, §3.4] goes as follows. Let $I = \mathbf{m}$ and let $J = \mathbf{n}$. For each $i \in I$ define $J_i = \{j \in J : p_{i,j} \neq 0\}$, and for each $j \in J$ define $I_j = \{i \in I : p_{i,j} \neq 0\}$. Then define equivalence relations

$$\mathscr{E}_I = \{(i_1, i_2) \in I \times I : J_{i_1} = J_{i_2}\},$$

and

$$\mathscr{E}_J = \{(j_1, j_2) \in J \times J : I_{j_1} = I_{j_2}\}.$$

Essentially, i_1, i_2 are $\mathscr{E}_I$ related if and only if row i_1 and row i_2 have 0's in exactly the same columns, and j_1, j_2 are $\mathscr{E}_J$ related if and only if column j_1 and column j_2 have 0's in exactly the same rows. Denote the $\mathscr{E}_I$-class containing i by $[i]_{\mathscr{E}_I}$ and the $\mathscr{E}_J$-class containing j by $[j]_{\mathscr{E}_J}$. The definition of these equivalences ensures that either $p_{x,y} = 0$ for all $(x,y) \in [i]_{\mathscr{E}_I} \times [j]_{\mathscr{E}_J}$, or $p_{x,y} \neq 0$ for all $(x,y) \in [i]_{\mathscr{E}_I} \times [j]_{\mathscr{E}_J}$ - we call this a non-zero block. For regular matrices there must be at least one non-zero block $[i]_{\mathscr{E}_I} \times [j]_{\mathscr{E}_J}$ for each $\mathscr{E}_I$-class $[i]_{\mathscr{E}_I}$. Similarly, for regular matrices there must be a non-zero block for each class $\mathscr{E}_J$-class. A matrix $(p_{i,j})$ is called normal if: (i) for every $\mathscr{E}_I$-class $[i]_{\mathscr{E}_I}$ there exists $j \in J$ such that $p_{i,j} = 1_G$ for all $i \in [i]_{\mathscr{E}_I}$, and (ii) for every $\mathscr{E}_J$-class $[j]_{\mathscr{E}_J}$ there exists $i \in I$ such that $p_{i,j} = 1_G$ for all $j \in [j]_{\mathscr{E}_J}$.

Now consider the following 4×4 matrix $(p_{i,j})$ with entries from a 0-group G^0 such that G is non-trivial:

$$(p_{i,j}) = \begin{pmatrix} 1_G & p_{1,2} & p_{1,3} & 0 \\ p_{2,1} & 1_G & 0 & p_{2,4} \\ p_{3,1} & 0 & 1_G & p_{3,4} \\ 0 & p_{4,2} & p_{4,3} & 1_G \end{pmatrix}.$$

Assume that the entries $p_{i,j}$ where $i \neq j$, or $i+j \neq 5$, that is to say those entries not on diagonals, are neither equal to 0 or equal to 1_G. The equivalence relations $\mathscr{E}_I$ and $\mathscr{E}_J$ are the diagonal relations on $\{1,2,3,4\}$ in this case, i.e. all their equivalence classes have size one. This matrix is considered to be normal by the standard definition. Define the matrix

$$(q_{i,j}) = (g_i^{-1} p_{i,j} u_j)$$

where

$$g_1 = 1_G, g_2 = p_{2,1}, g_3 = p_{3,1}, g_4 = p_{4,2} p_{1,2}^{-1},$$

and

$$u_1 = 1_G, u_2 = p_{1,2}^{-1}, g_3 = p_{1,3}^{-1}, g_4 = p_{2,4}^{-1} p_{2,1}.$$

Then we have that

$$(q_{i,j}) = \begin{pmatrix} 1_G & 1_G & 1_G & 0 \\ 1_G & g_2^{-1} u_2 & 0 & 1_G \\ 1_G & 0 & g_3^{-1} p_{3,3} u_3 & g_3^{-1} p_{3,4} u_4 \\ 0 & 1_G & g_4^{-1} p_{4,3} u_3 & g_4^{-1} p_{4,4} u_4 \end{pmatrix}$$

is another normal matrix in the same $G^m \times G^n$ orbit as $(p_{i,j})$ which has at least seven identity entries. The author argues that the aim of normalization is to guarantee as many identity entries as possible for a matrix when there non-zero entries are arbitrary, and that $(p_{i,j})$ does not achieve this goal. Therefore, the author believes that either $(p_{i,j})$ should not be considered normal, or that the property by which $(q_{i,j})$ is a superior normal matrix to $(p_{i,j})$ is part of the definition of normal matrices.

We will say that a subset T of $\mathbf{m} \times \mathbf{n}$ is a *normal type* if the bipartite graph $B(T)$ with vertex set

$$\{x_1, \ldots, x_m, y_1, \ldots, y_n\}$$

and edge set

$$\{(x_i, y_j) : (i,j) \in T\}$$

has no cycles. Let $(p_{i,j})$ be an $m \times n$ regular matrix with entries from a 0-group G^0 and let the subset T of $\mathbf{m} \times \mathbf{n}$ be a normal type. We will say that $(p_{i,j})$ is *T-normal* if and only if $p_{i,j} = 1_G$ for all $(i,j) \in T$, and we will call T the *normal type* of $(p_{i,j})$. Lemma 3.3.3 tells us precisely when there exists a matrix in a $G^m \times G^n$ orbit of $M_{m\times n}(G^0)$ of a given normal type.

Lemma 3.3.3. *Let $m,n > 0$ be integers. Let $(i_1,j_1),\ldots,(i_z,j_z) \in \mathbf{m} \times \mathbf{n}$. Then the following statements are equivalent:*

(i) *For any non-trivial group G and for any matrix $(p_{i,j}) \in M_{m\times n}(G^0)$ such that $p_{i_a,j_a} \neq 0$ for all $1 \leq a \leq z$, there exists another matrix $(q_{i,j})$ in the $G^m \times G^n$ orbit of $(p_{i,j})$ such that $q_{i_a,j_a} = 1_G$ for all $1 \leq a \leq z$,*

(ii) *The bipartite graph B with vertex set $\{x_1,\ldots,x_m,y_1,\ldots,y_n\}$ and edge set $\{(x_{i_a},y_{j_a}) : 1 \leq a \leq z\}$ has no cycles.*

Proof. ($\implies$) In order to prove (i) implies (ii) we assume (i) and not (ii) then show there is a contradiction. If (ii) does not hold then there exists $a_1,\ldots a_k$ such that

$$x_{a_0} \to y_{a_1} \to x_{a_2} \to \cdots \to y_{a_k} \to x_{a_0}$$

is a circuit of the graph B. Let G be a non-trivial finite group, and let g be a non-identity element of G. Let $(p_{i,j})$ be a matrix such that $p_{i_a,j_a} \neq 0$ for $1 \leq a \leq z$. Furthermore, assume that $(p_{i,j})$ satisfies:

$$g = \prod_{t\in\{0,2,\ldots,k-1\}} p_{i_{a_t},j_{a_{t+1}}} p^{-1}_{i_{a_{t+2}},j_{a_{t+1}}}, \tag{3.10}$$

which will later lead to the contradiction we seek. Note that we read subscripts of a modulo $k+1$. Such a $(p_{i,j})$ certainly exists, since we can choose any values for $p_{i_{a_0},j_{a_1}}, p_{i_{a_2},j_{a_3}},\ldots,p_{i_{a_{k-1}},j_{a_k}}$ and rearrange Equation 3.10 to deduce what $p_{i_{a_0},j_{a_k}}$ should be.

Statement (i) tells us that there exists $g = (g_1,\ldots,g_m) \in G^m$ and $u = (u_1,\ldots,u_m) \in G^n$ such that $(q_{i,j}) = (g_i^{-1} p_{i,j} u_j)$ satisfies $q_{i_a,j_a} = 1_G$ for $1 \leq a \leq z$. In particular, we have

$$g^{-1}_{i_\alpha} p_{i_\alpha,j_\alpha} u_{j_\alpha} = 1_G \qquad \text{for all } 1 \leq \alpha \leq z. \tag{3.11}$$

Herein, we will read subscripts of a modulo $k+1$. By combining the two equations from Equation 3.11 where α is such that (i_α,j_α) equals (a_t,a_{t+1}) or (a_t,a_{t+2}) we obtain

$$g^{-1}_{i_{a_t}} p_{i_{a_t},j_{a_{t+1}}} u_{j_{a_{t+1}}} = g^{-1}_{i_{a_{t+2}}} p_{i_{a_{t+2}},j_{a_{t+1}}} u_{j_{a_{t+1}}}, \tag{3.12}$$

which holds for all $0 \leq t \leq k$. Canceling the $u_{j_{a_{t+1}}}$ term and rearranging, we obtain:

$$g_{i_{a_t}} = p_{i_{a_t},j_{a_{t+1}}} p^{-1}_{i_{a_{t+2}},j_{a_{t+1}}} g_{i_{a_{t+2}}} \text{for all } 0 \leq t \leq k. \tag{3.13}$$

Repeated substitution of Equation 3.13 yields

$$\begin{aligned}
g_{i_{a_0}} &= p_{i_{a_0},j_{a_1}} p^{-1}_{i_{a_2},j_{a_1}} g_{i_{a_2}} \\
g_{i_{a_0}} &= p_{i_{a_0},j_{a_1}} p^{-1}_{i_{a_2},j_{a_1}} p_{i_{a_2},j_{a_3}} \left(p^{-1}_{i_{a_4},j_{a_3}} g_{i_{a_4}} \right) \\
&\vdots \\
g_{i_{a_0}} &= \left(\prod_{t \in \{0,2,\ldots,k-1\}} p_{i_{a_t},j_{a_{t+1}}} p^{-1}_{i_{a_{t+2}},j_{a_{t+1}}} \right) g_{i_{a_0}}.
\end{aligned}$$

This implies

$$1_G = \prod_{t \in \{0,2,\ldots,k-1\}} p_{i_{a_t},j_{a_{t+1}}} p^{-1}_{i_{a_{t+2}},j_{a_{t+1}}} \tag{3.14}$$

which contradicts Equation 3.10.

($\Longleftarrow$) Now we will show that (ii) implies (i). By (ii) there exists a bipartite graph B with vertices $\{x_1,\ldots,x_m,y_1,\ldots,y_n\}$ and edge set

$$\{(x_{i_a},y_{i_a}) : a \in \mathbf{z}\},$$

which has no cycles. In order to show (i), let G be any group, and let $(p_{i,j}) \in M_{m\times n}(G^0)$ be such that $p_{i_a,j_a} \neq 0$ for all $a \in \mathbf{z}$. Then we will show there exists $g = (g_1,\ldots,g_m) \in G^m$ and $u = (u_1,\ldots,u_m) \in G^n$ such that $(q_{i,j}) = (g_i^{-1} p_{i,j} u_j)$ satisfies $q_{i_a,j_a} = 1_G$ for $a \in \mathbf{z}$.

We will construct (g,u) via an iterative process. We define a series of elements of $G^m \times G^n$ starting with $(1_{G^m},1_{G^n})$ and finishing with (g,u). The series will be such that the action of one of the terms on $(p_{i,j})$ results in a matrix which has at least as many identity entries in the desired locations as the matrix produced by the action of a preceding term of the series.

In order to do this we need a walk[8] W_a for every connected component C_a of B, such that W_a includes every edge of the connected component C_a. Let there be $d+1$ connected components of B. Without loss of generality, for all $0 \leq s \leq d$ let

$$W_s = (x_{\alpha(s,1)},y_{\beta(s,1)}),(x_{\alpha(s,1)},y_{\beta(s,1)}),\ldots,(x_{\alpha(s,w_s)},y_{\beta(s,w_s)}),$$

[8] A walk on a graph is a sequence of edges, repeats are permitted.

denote a walk starting at $x_{\alpha(s,1)}$ and ending at $y_{\beta(s,w_s)}$ which includes all edges of the connected component C_a. Note we have implicitly defined w_s to be the length of the sequence W_s. Furthermore, it is implied that if t is odd then the walk W_s traverses the edge $(x_{\alpha(s,t)}, y_{\beta(s,t)})$ from $x_{\alpha(s,t)}$ to $y_{\beta(s,t)}$, and if t is even then the edge $(x_{\alpha(s,t)}, y_{\beta(s,t)})$ is traversed from $y_{\beta(s,t)}$ to $x_{\alpha(s,t)}$.

We now describe our iterative process for constructing (g,u). The sequence will contain the following terms in the order shown:

$$\begin{gathered}(g^{[0,0]},u^{[0,0]}),(g^{[0,1]},u^{[0,1]}),\ldots,(g^{[0,w_0]},u^{[0,w_0]}),\\(g^{[1,0]},u^{[1,0]}),(g^{[1,1]},u^{[1,1]}),\ldots,(g^{[1,w_1]},u^{[1,w_1]}),\\\vdots\\(g^{[d,0]},u^{[d,0]}),(g^{[d,1]},u^{[d,1]}),\ldots,(g^{[d,w_d]},u^{[d,w_d]}).\end{gathered}$$

The first term, $(g^{[0,0]},u^{[0,0]})$ will be equal to $(1_{G^m},1_{G^n})$ and the final term will be equal to the (g,u) we seek. For all terms, we will write

$$g^{[s,t]}=(g_1^{[s,t]},\ldots,g_m^{[s,t]}),$$

and

$$u^{[s,t]}=(u_1^{[s,t]},\ldots,u_n^{[s,t]}).$$

We will also define the matrices

$$(q_{i,j}^{[s,t]})=((g_i^{[s,t]})^{-1}\ p_{i,j}\ u_j^{[s,t]}) \tag{3.15}$$

for all $0\leq s\leq d$ and $0\leq t\leq w_s$ which will be used when defining the entries of our sequence of elements of $(g^{[s,t]},u^{[s,t]})$. If we can show that the matrix $(q_{i,j}^{[d,w_d]})$ satisfies

$$q_{i_a,j_a}^{[d,w_d]}=1_G$$

for all $a\in\mathbf{z}$ then we will be done, since this matrix is the result of the action of $(g^{[d,w_d]},u^{[d,w_d]})$ on $(p_{i,j})$. We define the sequence of elements of $G^m\times G^n$ now. Let

$$g_a^{[s,t]}=\begin{cases}g_a^{[s,t-1]} & \text{if t is odd, or } a\neq\alpha(s,t)\\ g_{\alpha(s,t)}^{[s,t-1]}q_{\alpha(s,t),\beta(s,t)}^{[s,t-1]} & \text{if } a=\alpha(s,t)\end{cases} \tag{3.16}$$

and

$$g^{[s+1,0]}=g^{[s,w_s]} \tag{3.17}$$

define the terms $g^{[s,t]}$. Essentially $g^{[s,t]}$ acts like the preceding term $g^{[s,t-1]}$ followed by left multiplying row $\alpha(s,t)$ by $(q^{[s,t-1]}_{\alpha(s,t),\beta(s,t)})^{-1}$. Let

$$u_a^{[s,t]} = \begin{cases} u_a^{[s,t-1]} & \text{if t is even, or } a \neq \beta(s,t) \\ u_{\beta(s,t)}^{[s,t-1]}(q^{[s,t-1]}_{\alpha(s,t),\beta(s,t)})^{-1} & \text{if } a = \beta(s,t) \end{cases} \tag{3.18}$$

and

$$u^{[s+1,0]} = u^{[s,w_s]} \tag{3.19}$$

define the terms $u^{[s,t]}$. Essentially $u^{[s,t]}$ acts like the preceding term $u^{[s,t-1]}$ followed by right multiplying column $\beta(s,t)$ by $(q^{[s,t-1]}_{\alpha(s,t),\beta(s,t)})^{-1}$.

We recall that the edge set of B is equal to

$$\{(x_{i_a} : y_{j_a}) : a \in \mathbf{z}\},$$

and so every edge of every walk $W_0, \ldots, W_d$ is of this form. Let W_s be the walk containing the edge (x_{i_a}, y_{i_a}) and let $0 \leq t \leq w_s$ be such that the tth term $(x_{\alpha(s,t)}, y_{\beta(s,t)})$ of W_s equals (x_{i_a}, y_{i_a}). If $t > 0$ is even then (3.16) tells us that

$$g_{i_a}^{[s,t]} = g_{i_a}^{[s,t-1]} q_{i_a,j_a}^{[s,t-1]},$$

and (3.18) tells us

$$u_{j_a}^{[s,t]} = u_{j_a}^{[s,t-1]}.$$

Combined with the definitions of $(q_{i,j}^{[s,t]})$ and $(q_{i,j}^{[s,t-1]})$, see (3.15), we deduce:

$$\begin{aligned} q_{i_a,j_a}^{[s,t]} &= (g_{i_a}^{[s,t]})^{-1} p_{i_a,j_a} u_{j_a}^{[s,t]} \\ &= (g_{i_a}^{[s,t-1]} q_{i_a,j_a}^{[s,t-1]})^{-1} p_{i_a,j_a} u_{j_a}^{[s,t-1]} \\ &= (q_{i_a,j_a}^{[s,t-1]})^{-1} (g_{i_a}^{[s,t-1]})^{-1} p_{i_a,j_a} u_{j_a}^{[s,t-1]} \\ &= (q_{i_a,j_a}^{[s,t-1]})^{-1} q_{i_a,j_a}^{[s,t-1]} \\ &= 1_G. \end{aligned}$$

Otherwise, if t is odd then (3.18) tells us that

$$u_{j_a}^{[s,t]} = u_{j_a}^{[s,t-1]} (q_{i_a,j_a}^{[s,t-1]})^{-1},$$

and (3.16) tells us

$$g_{i_a}^{[s,t]} = g_{i_a}^{[s,t-1]}.$$

Combined with the definitions of $(q_{i,j}^{[s,t]})$ and $(q_{i,j}^{[s,t-1]})$, see (3.15), we deduce:

$$\begin{aligned} q_{i_a,j_a}^{[s,t]} &= (g_{i_a}^{[s,t]})^{-1} p_{i_a,j_a} u_{j_a}^{[s,t]} \\ &= (g_{i_a}^{[s,t-1]})^{-1} p_{i_a,j_a} u_{j_a}^{[s,t-1]} (q_{i_a,j_a}^{[s,t-1]})^{-1} \\ &= q_{i_a,j_a}^{[s,t-1]} (q_{i_a,j_a}^{[s,t-1]})^{-1} \\ &= 1_G. \end{aligned}$$

We have now shown that if the tth term of W_s is equal to (x_{i_a}, y_{j_a}) then $q_{i_a,j_a}^{[s,t]}$ is equal to 1_G. Let us denote by $E_{s,t}$ the set

$$\{(x_{\alpha(s',t')}, y_{\beta(s',t')}) : 0 \leq s' < s,\ 0 \leq t' \leq w_s\} \cup \{(x_{\alpha(s,t')}, y_{\beta(s,t')}) : 0 \leq t' \leq t\}$$

containing all edges traversed by the walks $W_{s'}$ for $s' < s$ and the first t terms of W_s. Since B has no cycles, each connected component of B is a tree. Therefore, when the edge $(x_{\alpha(s,t)}, y_{\beta(s,t)})$ occurs during the walk W_s either: the path is visiting the destination vertex for the first time, or the path has already traversed this edge. In the former situation, we can say that $E_{s,t-1}$ contains no edges which are incident[9] on $x_{\alpha(s,t)}$ when $t > 0$ is even, or incident on $y_{\beta(s,t)}$ when t is odd. Therefore, if

$$q_{i_a,j_a}^{[s,t-1]} = 1_G \text{ for all } (i_a, j_a) \in E_{s,t-1}$$

then

$$q_{i_a,j_a}^{[s,t]} = 1_G \text{ for all } (i_a, j_a) \in E_{s,t-1}.$$

This is because there is no $(i_a, j_a) \in E_{s,t-1}$ such that $i_a = \alpha(s,t)$ when $t > 0$ is even, or $j_a = \beta(s,t)$ when t is odd. By induction, we can deduce that the statement: $q_{i_a,j_a}^{[s,t]} = 1_G$ for all $(i_a, j_a) \in E_{s,t}$, holds for all $0 \leq s \leq d$ and $0 \leq t \leq w_s$. Finally, since the every edge

$$\{(x_{i_a}, y_{j_a}) : a \in \mathbf{z}\}$$

of B is traversed by one of the walks W_s, and if the tth term of W_s is equal to (x_{i_a}, y_{j_a}) then $q_{i_a,j_a}^{[s,t]} = 1_G$, we deduce that $(q_{i,j})^{[d,w_d]}$ satisfies

$$(q_{i_a,j_a})^{[d,w_d]} = 1_G$$

[9] An edge (a,b) is said to be incident on a and on b, the two vertices it connects.

for all $a \in \mathbf{z}$. Recall that $(q_{i_a,j_a})^{[d,w_d]}$ is the result of acting on $(p_{i,j})$ by $(g^{[d,w_d]}, u^{[d,w_d]}) \in G^m \times G^n$, so we have found an element of the orbit $G^m \times G^n$ orbit of $(p_{i,j})$ which satisfies the required condition. □

Recall that in Section 3.2 we defined $M_{m\times n}(G^0, S)$ to be the subset of $M_{m\times n}(G^0)$ containing the matrices which have binary shape equal to S. Let $T \subseteq \mathbf{m} \times \mathbf{n}$ be a normal type and let $S = (s_{i,j})$ be an $m \times n$ binary matrix satisfying $s_{i,j} \neq 0$ for all $(i,j) \in T$. Then Lemma 3.3.3 tells us that for all matrices in $M_{m\times n}(G^0, S)$ there exists a T-normal matrix in the same $G^m \times G^n$ orbit. We will call the process of finding a T-normal matrix of an orbit of $G^m \times G^n$ *T-normalization*. It is clear that T-normalization is only possible when the binary shape of matrices in a given orbit is *compatible* with T, in the sense that they don't have 0 entries with indices corresponding to elements of T.

As mentioned earlier, we want a method of normalization which guarantees as many identity entries as possible. Thus we want to use only the largest normal types T, which will depend on the binary shape of the matrix we want to normalize. If S is a $m \times n$ regular binary matrix and $T \subseteq \mathbf{m} \times \mathbf{n}$ is a normal type compatible with S then we will say T is *S-maximal* if T has the largest possible cardinallity amongst normal types compatible with S. Proposition 3.3.4 determines the maximum size of a normal type compatible with a binary shape.

Proposition 3.3.4. *Let $m,n > 0$ be integers. Let G be a finite group. Let $S = (s_{i,j})$ be a $m \times n$ regular binary matrix. Then the size of a S-maximal normal type T is $m+n-k$, where k is the number of connected components of the bipartite graph B with vertex set*

$$\{x_1, \ldots, x_m, y_1, \ldots, y_n\}$$

and edge set

$$\{(x_i, y_j) : s_{i,j} \neq 0\}.$$

Proof. In order for $T \subseteq \mathbf{m} \times \mathbf{n}$ to be a normal type compatible with S the set of edges

$$E_T = \{(x_i, y_j) : (i,j) \in T\}$$

must be a subset of the edges of B and the subgraph of B with the same vertex set and edge set E_T must have so cycles. Every connected component of a graph with no cycles is a tree. A tree has one less edge than vertices. Therefore, a graph where every connected component is a tree has $m+n$ minus the number of connected components edges. The least number of connected components of a subgraph of B with no cycles is the same as the number of connected components of B. Thus the size of a normal type compatible with S is at most $m+n-k$. It should be clear that a normal type of this size does exist, finding one simply involves choosing

a spanning tree[10] for each connected component of $B(S)$. The union of the edges of these trees corresponds to a S-maximal normal type with size $m+n-k$. □

We now present some examples to illustrate our theory of normal matrices and normal types. Herein, when S is a binary matrix we will define the bipartite graph $B(S)$, as seen in Proposition 3.3.4, to be the one with vertex set

$$\{x_1,\ldots,x_m,y_1,\ldots,y_n\}$$

and edge set

$$\{(x_i,y_j) : s_{i,j} \neq 0\}.$$

Our first example has $B(S)$ being a connected graph.

Example 3.3.5. Let $P = (p_{i,j})$ be a 4×4 matrix with entries from a 0-group G^0 of the form:

$$(p_{i,j}) = \begin{pmatrix} p_{1,1} & p_{1,2} & 0 & 0 \\ p_{2,1} & p_{2,2} & 0 & p_{2,4} \\ 0 & 0 & p_{3,3} & p_{3,4} \\ 0 & 0 & p_{4,3} & p_{4,4} \end{pmatrix}.$$

Then Figure 3.1 shows the bipartite graph $B(S)$ associated with the binary shape S of P. A subset of the edges of $B(S)$ which forms a tree has been highlighted in red. If we set

$$T = \{(1,1),(1,2),(2,1),(2,4),(3,4),(4,3),(4,4)\}$$

then T is a normal type compatible with S, and the highlighted subgraph is the bipartite graph $B(T)$ associated with T. Moreover T is a S-maximal normal type, since $B(S)$ is connected and T has size $7 = 4+4-1$. A T-normalization of P would have the following form:

$$\begin{pmatrix} 1_G & 1_G & 0 & 0 \\ 1_G & a & 0 & 1_G \\ 0 & 0 & b & 1_G \\ 0 & 0 & 1_G & 1_G \end{pmatrix}$$

form some $a,b \in G$.

The next example is such that the bipartite graph associated with the binary shape has two connected components.

[10] A spanning tree is a subgraph which is a tree and includes every vertex.

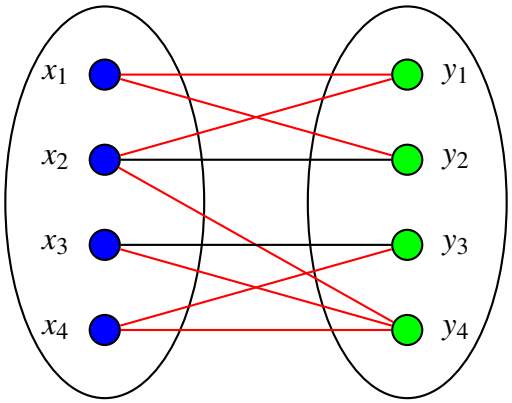


Fig. 3.1 The bipartite graph $B(S)$ with the edges of the subgraph $B(T)$ highlighted in red, from Example 3.3.5

Example 3.3.6. Let $P = (p_{i,j})$ be a 5×5 matrix with entries from a 0-group G^0 of the form:

$$(p_{i,j}) = \begin{pmatrix} p_{1,1} & p_{1,2} & 0 & 0 & 0 \\ p_{2,1} & p_{2,2} & 0 & 0 & 0 \\ 0 & 0 & p_{3,3} & p_{3,4} & p_{3,5} \\ 0 & 0 & p_{4,3} & 0 & p_{4,5} \\ 0 & 0 & p_{5,3} & p_{5,4} & 0 \end{pmatrix}$$

Then Figure 3.2 shows the bipartite graph $B(S)$ associated with the binary shape S of P. A subset of the edges of $B(S)$ which forms a graph with no cycles has been highlighted in red. If we set

$$T = \{(1,1),(1,2),(2,1),(3,3),(3,4),(3,5),(4,3),(5,3)\}$$

then T is a normal type compatible with S, and the highlighted subgraph is the bipartite graph $B(T)$ associated with T. Moreover T is a S-maximal normal type, since $B(S)$ has two connected components and T has size $8 = 5+5-2$. A T-normalization of P would have the following form:

$$(p_{i,j}) = \begin{pmatrix} 1_G & 1_G & 0 & 0 & 0 \\ 1_G & p_{2,2} & 0 & 0 & 0 \\ 0 & 0 & 1_G & 1_G & 1_G \\ 0 & 0 & 1_G & 0 & a \\ 0 & 0 & 1_G & b & 0 \end{pmatrix}$$

form some $a, b \in G$.

The normalization theorem in [19, Theorem 3.4.2], which stated that for every matrix $P \in M_{m\times n}(G^0)$ there is another matrix Q which is normal such that $\mathscr{M}^0[G;P]$ is isomorphic

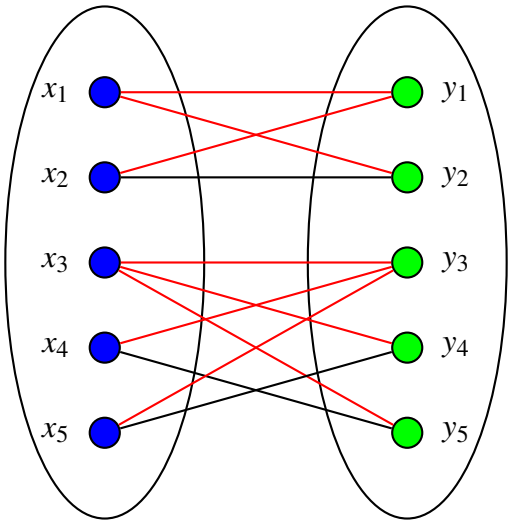


Fig. 3.2 The bipartite graph $B(S)$ with the edges of the subgraph $B(T)$ highlighted in red, from Example 3.3.6

to $\mathscr{M}^0[G;Q]$. The following result is our adaption of that theorem to our definition of normal matrices.

Theorem 3.3.7. *Let $P = (p_{i,j})$ be a $m \times n$ regular matrix with entries from G^0. Let S denote the binary shape of P. Then there exists an S-maximal normal type T and a T-normal matrix Q such that $\mathscr{M}^0[G;P]$ is isomorphic to $\mathscr{M}^0[G;Q]$.*

Proof. By Proposition 3.3.4 there exists a S-maximal normal type T of size $m+n-k$ where k is the number of connected components of $B(S)$. Lemma 3.3.3 tells us that there is a T-normal matrix Q in the same $G^m \times G^n$ orbit as P. Two matrices being in the same $G^m \times G^n$ orbit is a stronger property than those matrices being in the same $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbit. For P and Q, the latter statement is true if and only if $\mathscr{M}^0[G;P]$ is isomorphic to $\mathscr{M}^0[G;Q]$. □

We can apply this result to reducing the number of matrices which CANONICALIMAGE must be applied to when enumerating a transversal of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbits of $M_{m\times n}(G^0)$. We continue utilising the strategy of applying CANONICALIMAGE to the sets $M_{m\times n}(G^0,S)$ for S in $\mathbb{S}$, a transversal of the $S_m \times S_n$ orbits of $m \times n$ regular binary matrices. Theorem 3.3.7 tells us that every orbit of $M_{m\times n}(G^0,S)$ contains a matrix with S-maximal normal type. Therefore we will still find a transversal of $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbits of $M_{m\times n}(G^0)$ if for each $S \in \mathbb{S}$ we choose a S-maximal normal type T and we only apply CANONICALIMAGE to T-normal matrices in $M_{m\times n}(G^0,S)$.

Let S be a $m \times n$ binary shape and let T be a S-maximal normal type. Then it follows from Theorem 3.3.7 and Proposition 3.2.1 that the set $M_{m\times n}(G^0, S, T)$

$$\{(p_{i,j})_{i,j} \in M_{m\times n}(G^0, S) : (p_{i,j}) \text{ is } T\text{-normal}\} \tag{3.20}$$

contains representatives of all $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbits of $M_{m\times n}(G^0)$ which intersect $M_{m\times n}(G^0, S)$. We then have the following theorem.

Theorem 3.3.8. *Let $\mathbb{S}$ be a transversal of the $S_m \times S_n$ orbits of regular $m \times n$ binary matrices. For each S in $\mathbb{S}$ let T_S be a S-maximal normal type. Then the set*

$$\bigcup_{S\in\mathbb{S}} M_{m\times n}(G^0, S, T_S)$$

contains representatives of every $(G^m \times G^n) \rtimes (S_m \times S_n \times \mathrm{Aut}(G^0))$ orbit of $M_{m\times n}(G^0)$.

We conclude this section with an example showing how Theorem 3.3.8 in reduces the number of matrices we apply CANONICALMATRIX to.

Example 3.3.9. Consider a non-trivial group G and the 3×3 binary shape:

$$S := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}. \tag{3.21}$$

Let us define two normal types compatible with S:

$$T_1 = \{(1,1),(1,2),(1,3),(2,1),(3,1)\},$$

and

$$T_2 = \{(1,1),(3,2),(2,3)\}.$$

The normal type T_1 is S-maximal, with elements 5, whereas the normal type T_2 has only 3 elements. Note that a T_2-normal matrix would be considered a normal matrix by the standard definition. The set $M_{m\times n}(G, S, T_1)$ contains $|G|^2$ matrices:

$$M_{m\times n}(G, S, T_1) = \left\{ \begin{pmatrix} 1_G & 1_G & 1_G \\ 1_G & 0 & a \\ 1_G & b & 0 \end{pmatrix} : a, b \in G \right\}, \tag{3.22}$$

where as the set $M_{m\times n}(G,S,T_1)$ contains $|G|^4$ matrices:

$$M_{m\times n}(G,S,T_1)=\left\{\begin{pmatrix}1_G & a & b\\ c & 0 & 1_G\\ d & 1_G & 0\end{pmatrix} : a,b,c,d\in G\right\}. \tag{3.23}$$

In Section 3.2 we showed that we could find a transversal of the $(G^m\times G^n)\rtimes(S_m\times S_n\times \mathrm{Aut}(G^0))$ orbits of $M_{m\times n}(G^0)$ by applying CANONICALIMAGE to all matrices in

$$\bigcup_{S\in\mathbb{S}} M_{m\times n}(G^0,S)$$

where $\mathbb{S}$ is a transversal of $S_m\times S_n$ orbits of $m\times n$ binary matrices. This requires $|\mathbb{S}||G|^s$ applications of CANONICALIMAGE, where s is the number of 1's in the binary matrix S. Applying CANONICALIMAGE to all matrices in

$$\bigcup_{S\in\mathbb{S}} M_{m\times n}(G^0,S,T_2)$$

requires $|\mathbb{S}||G|^4$ applications of CANONICALIMAGE. Finally, applying CANONICALIMAGE to all matrices in

$$\bigcup_{S\in\mathbb{S}} M_{m\times n}(G^0,S,T_1)$$

requires $|\mathbb{S}||G|^2$ applications of CANONICALIMAGE.

3.4 Results

The GitHub repository https://github.com/ChristopherRussell/0-simple-semigroups contains **GAP** code relating to this chapter, written by the author. The main method is called ALLZEROSIMPLESEMIGROUPS and can take either two or four arguments. The two argument method takes an order k and a boolean, then returns a list of representatives of the isomorphism classes of 0-simple semigroups of order k. Depending on the boolean value, these representatives can be found up to anti-isomorphism and isomorphism. The four argument method takes a group G, positive integers $m,n>0$, and a boolean, and returns a list of representatives of the isomorphism classes of 0-simple semigroups constructed from matrices in $M_{m\times n}(G^0)$. Again, the boolean value can be used to find representatives up to anti-isomorphism and isomorphism. The repository does not contain a database constructed using these methods, rather the user can construct the cases they are interested in.

The authors code was able to construct all 0-simple semigroups of order at most 49 up to isomorphism. We compared the number of semigroups constructed to the result of our enumeration in Section 2.7 and the numbers agreed in all cases, which is strong evidence for their correctness as these numbers were found by completely different methods. The limiting factor for constructing all 0-simple semigroups up to isomorphism up to some order k is the square or roughly square[11] binary matrices, which correspond to $\mathscr{H}$-trivial 0-simple semigroups. As seen in Section 2.7 when we consider all isomorphism classes of 0-simple semigroups of order less than some integer k, the vast majority are $\mathscr{H}$-trivial and the greater k is the larger the majority tends to be. Our algorithms were unable to construct all 0-simple semigroups of order 50, in particular the case corresponding to 7×7 binary matrices. In Section 2.7 we calculated that there are $26,610,810$ 0-simple semigroups of order 50 up to isomorphism, and all but 4 of these are $\mathscr{H}$-trivial.

The benefits of the theory of binary shapes and normalization, which were our main insights in this chapter, are only useful for cases where the 0-simple semigroups are not $\mathscr{H}$-trivial. Therefore our database should not be evaluated solely on being complete up to order 49. Rather, we should evaluate our algorithms on for how many parameters G, m, n it can find a complete list of representatives for the isomorphism classes of 0-simple semigroups constructed from matrices in $M_{m\times n}(G^0)$. Our code can tackle these cases for many G, m, n such that $m * n * |G| + 1 > 49$.

When applied our code to construct all 0-simple semigroups of order less than 50 up to isomorphism most calculations were possible using a single process of **GAP** on a 2016 MacBook Pro with 2.6GHz quad-core Intel Core i7, 16GB of 2133MHz LPDDR3 RAM. However a few of the hardest calculations were performed in parallel with computers owned by CIRCA in the School of Computer Science at St Andrews. They are Lovelace, Babbage: 64 cores (Bulldozer), 512GB RAM each. Mandel, Kovacs: 4 cores, 8 threads, 64GB RAM each, Xeon E3 machines. Cormac: 20 cores, 40 threads, 128GB RAM, this is a 2 socket Xeon E5 machine. Parallel computation was only used to find transversals of the $S_m \times S_n$ orbits of 6×6, 6×7, and 6×8 binary matrices.

[11] By roughly square we mean a $m \times n$ matrix where $|m - n|$ is small, e.g. 1 or 2.

Chapter 4

Counting congruence free semigroups

4.1 Introduction

Congruence free semigroups are to semigroup congruences what simple groups are to normal subgroups. In this sense, congruence free semigroups are the semigroup analogy of simple groups. With simple groups having been such a lively area of research in recent times, for example the classification of finite simple groups [11] up to isomorphism is one of the most impressive results in modern mathematics, it is natural to be interested in their semigroup counterparts. A classification of congruence free semigroups was covered earlier in Theorem 1.7.1.

Our interest lies in counting the number of these semigroups up to isomorphism for each order. This bears similarity to the long active problem of counting finite groups by order, which has had many contributors over many years. The groups of order at most 6 were counted in 1854 by Cayley [7], then many more hand calculated results followed before computational methods took over and eventually extended our knowledge to order 2000 [2] and beyond. In the case of finite simple groups, counting by order is not particularly interesting. Most orders have zero instances and the rest have one or two (with the latter situation occurring only at order 20160 [38] or in a certain infinite family of cases [21]). The counting of finite semigroups up to isomorphism by order has also attracted academic interested in recent times [8, 9, 13, 31]. Counting finite congruence free semigroups up to isomorphism appears to be a relatively unbroached subject for all but the smallest values.

Theorem 1.7.1 states that there are three possibilities for congruence free semigroups. First, all finite simple groups are congruence free semigroups, in this case we offer no new results. Second, there are six semigroups of order at most two up to isomorphism, and they are all congruence free. Finally, a finite congruence free semigroup which is not of the two types just described must be a 0-simple semigroup. Furthermore if $\mathscr{M}^0[G;I,\Lambda;P]$ is a congruence free Rees 0-matrix semigroup, then the group G is the trivial group and P is regular (no rows or

columns are all-zero) with all rows unique and all columns unique. In this chapter, we will count the isomorphism classes of 0-simple semigroups of this last type. This problem boils down to counting binary matrices with all rows and columns distinct, up to permutations of the rows and columns independently.

The enumeration of various classes of binary matrices has been an area interest to many researchers. At least in part, binary matrices are often of coincidental interest due to their correspondence with numerous other mathematical objects, especially in graph theory as adjacency matrices, for example graphs [14] or hypergraphs [28]. Furthermore equivalence classes of binary matrices correspond with interesting classes of the related objects. For example, when the $n \times n$ binary matrices are viewed as adjacency matrices of directed graphs of degree n then the orbits of the action of S_n by (permuting the rows and columns simultaneously) on this set are in correspondence with the isomorphism classes of the corresponding graphs. Harrison [16] studied binary matrices up to row and column permutations, and column complementation, noting applications in switching theory. Incidence matrices (binary matrices without zero rows or columns) are another class of binary matrices which have been enumerated subject to having various additional properties, and up to row and column permutation [4, 33].

Perhaps the closest work to the subject of this chapter was by Houghton [17] who described formulae for counting 0-simple semigroups up to isomorphism. The author's own work in this area is described in Chapter 2. In the context of 0-simple semigroups, the constraints of being congruence free has allowed for a greatly superior method for counting.

Despite the best efforts of the author, this chapter is heavy in notation. To help there is a list of the symbols used at the end of the Chapter which can be used as a reference, see Chapter 4 Symbols.

4.2 Counting orbits

Herein we will frequently consider $m \times n$ binary matrices as functions from $\mathbf{m} \times \mathbf{n}$ to $\{0,1\}$. We will denote the subset of $\{0,1\}^{\mathbf{m}\times\mathbf{n}}$ of matrices with all rows distinct and all columns distinct by $X_{m,n}$. The natural action of $(\rho,\sigma) \in S_m \times S_n$ on $f \in \{0,1\}^{\mathbf{m}\times\mathbf{n}}$ by permuting rows and columns is

$$(x,y)f^{(\rho,\sigma)} = (x\rho^{-1}, y\sigma^{-1})f.$$

Note that $X_{m,n}$ is invariant under this action. Our aim is to count the number of orbits of $X_{m,n}$ under this action, which correspond to the isomorphism classes of congruence free semigroups.

To do this we will use the orbit-counting theorem:

$$|X_{m,n}//S_m \times S_n| = \frac{1}{m!\,n!} \sum_{(\rho,\sigma)\in S_m\times S_n} |X_{m,n}^{(\rho,\sigma)}|$$

where $A//B$ denotes the orbits of B in its action on A and $X_{m,n}^{(\rho,\sigma)}$ denotes the elements of $X_{m,n}$ fixed by (ρ,σ). To use this theorem, we need a formula for $|X_{m,n}^{(\rho,\sigma)}|$ for each (ρ,σ) in $S_m \times S_n$. However we can make life easier for ourselves by using the fact that conjugate elements fix the same number of elements. Therefore we only need to determine $|X_{m,n}^{(\rho,\sigma)}|$ for one representative (ρ,σ) of each conjugacy class $[(\rho,\sigma)]$ of $S_m \times S_n$ and multiply by the size of the conjugacy class. The conjugacy classes of the symmetric group correspond to cycle type and the conjugacy classes of a direct product are the Cartesian products of the conjugacy classes of the factors. The cycle type of a permutation ρ in S_m can be described by a tuple $(i_1,\ldots,i_m)$ where the sum of entries is equal to m and the entry i_j denotes the number of j-cycles of ρ. Using our knowledge of conjugacy classes we may restate the number of orbits of $X_{m,n}$ as:

$$|X_{m,n}//S_m \times S_n| = \frac{1}{m!\,n!} \sum_{i_1+2i_2+\cdots+mi_m=m} \frac{m!}{\prod_{a=1}^m a^{i_a} i_a!} \sum_{j_1+2j_2+\cdots+nj_n=n} \frac{n!}{\prod_{b=1}^n b^{j_b} j_b!} |X_{m,n}^{(\rho_i,\sigma_j)}|$$

where ρ_i denote a cycle of type $(i_1,\ldots,i_m)$ and σ_j denotes a cycle of type $(j_1,\ldots,j_n)$. This expression may be simplified to

$$|X_{m,n}//S_m \times S_n| = \sum_{i_1+2i_2+\cdots+mi_m=m} \sum_{j_1+2j_2+\cdots+nj_n=n} \frac{|X_{m,n}^{(\rho,\sigma)}|}{\prod_{a=1}^m a^{i_a} i_a! \prod_{b=1}^n b^{j_b} j_b!}$$

where $i_1,\ldots,i_m >= 0$ satisfying $i_1+2i_2+\cdots+mi_m = m$ defines a cycle type of a permutation in S_m and

$$\frac{m!}{\prod_{a=1}^m a^{i_a} i_a!}$$

is the size of the conjugacy class of elements with this cycle type. Our next task is to determine $|X_{m,n}^{(\rho,\sigma)}|$ for a representative (ρ,σ) of each conjugacy class of $S_m \times S_n$.

4.3 Matrices fixed by a pair of permutations

We first examine which matrices in $\{0,1\}^{\mathbf{m}\times\mathbf{n}}$ are fixed by a pair of permutations (ρ,σ) in $S_m \times S_n$. This is an step towards our ultimate goal of determining which of these matrices also have distinct rows and columns. We will begin by defining the equivalence relation $R_{\rho,\sigma}$ on

$\mathbf{m} \times \mathbf{n}$ to be the one with classes corresponding to the orbits of the subgroup $\langle(\rho,\sigma)\rangle$ of $S_m \times S_n$ generated by (ρ,σ). Then we can describe exactly when a matrix will be fixed by (ρ,σ) with the following lemma.

Lemma 4.3.1. *A $m \times n$ binary matrix $f \in \{0,1\}^{\mathbf{m}\times\mathbf{n}}$ is fixed by $(\rho,\sigma) \in S_m \times S_n$ if and only if*

$$(x_1,y_1)R_{\rho,\sigma}(x_2,y_2) \implies (x_1,y_1)f = (x_2,y_2)f$$

for all $(x_1,y_1),(x_2,y_2) \in \mathbf{m}\times\mathbf{n}$. In other words, for all (x,y) in a class of $R_{\rho,\sigma}$ the corresponding entries $(x,y)f$ of f have the same value.

Proof. Let $f \in \{0,1\}^{\mathbf{m}\times\mathbf{n}}$ be fixed by $(\rho,\sigma) \in S_m \times S_n$. This equivalent to the statement: $(x,y)f = (x\rho,y\sigma)f$ for all $(x,y) \in \mathbf{m}\times\mathbf{n}$. In turn, the statement $(x,y)f = (x\rho,y\sigma)f$ for all $(x,y) \in \mathbf{m}\times\mathbf{n}$ is true if and only if

$$(x,y)f = (x\rho,y\sigma)f = (x\rho^2,y\sigma^2)f = \cdots = (x\rho^i,y\sigma^i)f$$

for all $(x,y) \in \mathbf{m}\times\mathbf{n}$ and all $i \in \mathbb{N}$. Finally, this is equivalent to the statement: $(x_1,y_1)f = (x_2,y_2)f$ for any pairs $(x_1,y_1),(x_2,y_2)$ which are in the same orbit of $\langle(\rho,\sigma)\rangle$. □

We will identify equivalence relations with their corresponding partitions[1], so that if R is an equivalence relation we will also write R to denote the related partition. We can see that matrices in $\{0,1\}^{\mathbf{m}\times\mathbf{n}}$ fixed by $(\rho,\sigma) \in S_m \times S_n$ correspond to functions $f : R_{\rho,\sigma} \to \{0,1\}$ from the partition $R_{\rho,\sigma}$ to $\{0,1\}$ since Lemma 4.3.1 tells us that entries with indices from the same class must have the same value. For such an f we will write f' to denote the corresponding function from $\mathbf{m}\times\mathbf{n}$ to $\{0,1\}$, that is to say we set

$$(x,y)f' = [(x,y)]_{R_{\rho,\sigma}}f$$

for all $(x,y) \in \mathbf{m}\times\mathbf{n}$ where $[(x,y)]_{R_{\rho,\sigma}}$ refers to the class of (x,y) in $R_{\rho,\sigma}$.

Remark 4.3.2. *Consider an arbitrary binary matrix fixed by $((1\,2\,\ldots\,m),(1\,2\,\ldots\,n))$.*

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a_{m-1\,n} \\ a_{m1} & \cdots & \cdots & a_{m\,n-1} & a_{mn} \end{pmatrix}$$

[1] The equivalence classes of an equivalence relation form a partition.

The permutation $((1\,2\,\dots\,m),(1\,2\,\dots\,n))$ *permutes its entries by:*

$$a_{i\,j} \to a_{i+1\,j+1} \to \cdots \to a_{i-1\,j-1} \to a_{i\,j}$$

where indices are taken mod m and n, and 0 *mod x is replaced by x for* $x = m,n$*. If a matrix is fixed by this permutation then all the entries in this orbit must be the same. There are 2 to the power of the number of orbits (choose 0 or 1 for each orbit).*

Remark 4.3.3. *The number of orbits of the group generated by* $\langle(1\,2\,\dots\,m),(1\,2\,\dots\,n)\rangle$ *acting on* $\{1,2,\dots,m\}\times\{1,2,\dots,n\}$ *is equal to the greatest common divisor*[2] *of m and n. Each orbit has the same size, which is equal to the least common multiple of m and n. As mentioned earlier, entries of a matrix fixed by* $(1\,2\,\dots\,m),(1\,2\,\dots\,n)$ *with indices in the same orbit of this action must be equal. Examples of matrices of sizes* 4×4*,* 3×4 *and* 4×6 *are shown below with the entries corresponding to orbits highlighted, with distinct colours indicating different orbits.*

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \quad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix} \quad \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{pmatrix}$$

Herein we will define dom to return the domain of a function, including permutations, for example if $\rho \in S_m$ then $\mathrm{dom}(\rho)$ will refer to $\{1,\dots,m\}$. However, in the case that ρ is a permutation composed of disjoint cycles $\rho_1,\dots,\rho_r$ where ρ_i is the cycle $(\alpha_1\ \dots\ \alpha_z)$ we will write $\mathrm{dom}(\rho_i)$ to refer to the set of moved points $\{\alpha_1,\dots,\alpha_z\}$ of the disjoint cycle ρ_i of ρ. When ρ_i is a disjoint cycle of length 1 we define $\mathrm{dom}(\rho_i)$ to be the element of that cycle, even though that element is a fixed point not a moved point[3]. The following example shows how matrices fixed by a pair of non-cyclic permutations can be understood in terms of matrices fixed by a pair of cyclic permutations, such as those we saw in Example 4.3.3.

Remark 4.3.4. *Let* $\rho \in S_m$ *be composed of the cycles* $\rho_1,\dots,\rho_r$ *where. Without loss of generality, we may assume that:*

$$\mathrm{dom}(\rho_1) = \{1,\dots,|\rho_1|\},$$

[2] Let $\rho = (1\,2\,\dots\,m)$ and $\sigma = (1\,2\,\dots\,n)$. Notice that for any $(i,j) \in \mathbf{m}\times\mathbf{n}$ we have that $(i\rho^k, j\sigma^k) = (i,j)$ when k is divisible by the orders of ρ and σ. Therefore the length of the orbit of (i,j) in this example is the least common multiple of m and n. Since (i,j) was arbitrary in $m\times n$ all orbits have length $\mathrm{lcm}(m,n)$ and the total number of orbits is $mn/\mathrm{lcm}(m,n)$ which is equal to the greatest common divisor of m and n.

[3] The author acknowledges that their use of dom notation is potentially confusing abuse of notation. The disjoint cycles of a permutation $\rho \in S_m$ are really cyclic permutations in S_m, which have domain $\{1,\dots,m\}$ which is the same as $\mathrm{dom}(\rho)$. Defining the domain of the disjoint cycles $\rho_1,\dots,\rho_r$ of $\rho \in S_m$ to be proper subsets of $\{1,\dots,m\}$ also has the issue that the identity $\rho = \rho_1\cdots\rho_r$ becomes incorrect. However, as noted, referring to the moved points of a disjoint cycle is also problematic since we want to be able to refer to the elements of a disjoint cycle of length one.

$$\mathrm{dom}(\rho_2) = \{|\rho_1|+1,\ldots,|\rho_1|+|\rho_2|\},$$
$$\vdots$$
$$\mathrm{dom}(\rho_r) = \{m-|\rho_r|+1,\ldots,m\}.$$

In particular, this means that

$$\rho = (1\,2\ldots|\rho_1|)(|\rho_1|+1\ldots|\rho_1|+|\rho_2|)\cdots(m-|\rho_r|+1\ldots m)$$

i.e. every cycle of ρ can be written a sequence of consecutive integers. Similarly, Let $\sigma \in S_n$ be composed of the disjoint cycles $\sigma_1,\ldots,\sigma_s$ where

$$\mathrm{dom}(\sigma_1) = \{1,\ldots,|\sigma_1|\},$$
$$\mathrm{dom}(\sigma_2) = \{|\sigma_1|+1,\ldots,|\sigma_1|+|\sigma_2|\},$$
$$\vdots$$
$$\mathrm{dom}(\sigma_s) = \{m-|\sigma_s|+1,\ldots,n\}.$$

In particular, this means that

$$\sigma = (1\,2\ldots|\sigma_1|)(|\sigma_1|+1\ldots|\sigma_1|+|\sigma_2|)\cdots(n-|\sigma_s|+1\ldots n).$$

For all $(i,j) \in \mathbf{r}\times\mathbf{s}$, the sub-matrix[4] *$f_{i,j}$ on the domain $\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)$ is a matrix fixed by the $(\rho_i,\sigma_j) \in S_{\mathrm{dom}(\rho_i)}\times S_{\mathrm{dom}(\sigma_j)}$*[5]*. When we restrict our attention to this sub-matrix, the situation will look like what we saw in Example 4.3.2. That is to say, $f_{i,j}$ looks like a $|\rho_i|\times|\sigma_j|$ matrix fixed by the pair of cyclic permutations $(1\ldots|\rho_i|),(1\ldots|\sigma_j|))$. A matrix fixed by (ρ,σ) is composed of all the $r*s$ aforementioned sub-matrices fixed by the corresponding pairs of disjoint cycles. With the way we defined ρ and σ, we can roughly illustrate how the whole matrix is composed of these sub-matrices:*

$$\left(\begin{array}{c|c|c|c} f_{1,1} & f_{1,2} & \cdots & f_{1,s} \\ \hline f_{2,1} & f_{2,2} & \cdots & f_{2,s} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline f_{r,1} & f_{r,2} & \cdots & f_{r,s} \end{array}\right).$$

[4] The sub-matrix $f_{i,j}$ is the function from $\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)$ to $\{0,1\}$ such that $(x,y)f' = (x,y)f$ for all $(x,y) \in \mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)$.

[5] Here we are considering $\rho_i \in S_{\mathrm{dom}(\rho_i)}$ as the restriction of the function ρ to the domain (and range) $\mathrm{dom}(\rho_i)$, and σ_j is defined similarly.

We chose ρ and σ so that the most standard way to sketch this generic matrix fixed by (ρ,σ) would be such that all the $f_{i,j}$ sub-matrices have row and column domains consisting of consecutive integers. However, the intuition of this picture holds for matrices fixed by other pairs of permutations in $S_m \times S_n$, even if they are not of the special form that we have assumed of (ρ,σ). After all, a matrix is just a function from the cartesian product of the row indices and column indices to the set which the entries come from. When we draw a matrix we are choosing an ordering of the rows and an ordering of the columns, but these orderings are not part of the matrix definition.

The following proposition summarises the preceding remarks concerning the number of matrices fixed by a pair of permutations.

Proposition 4.3.5. *Let $\rho \in S_m$ and $\sigma \in S_n$ be permutations such that ρ is composed of the disjoint cycles $\rho_1,\ldots,\rho_r$ and σ is composed of the disjoint cycles $\sigma_1,\ldots,\sigma_s$. Then the number of $m \times n$ binary matrices fixed by (ρ,σ) is equal to*

$$\prod_{i=1}^{r}\prod_{j=1}^{s} 2^{gcd(|\mathrm{dom}(\rho_i)|,|\mathrm{dom}(\sigma_j)|)}.$$

Proof. By Lemma 4.3.1, the entries of a matrix fixed by (ρ,σ) with indices in the same orbit of the subgroup $\langle(\rho,\sigma)\rangle$ of $S_m \times S_n$ must be equal. Therefore the number of binary matrices fixed by (ρ,σ) is 2 to the power of the number of orbits of $\langle(\rho,\sigma)\rangle$. The number of orbits of $\langle(\rho,\sigma)\rangle$ is the product of the number of orbits of all the $\langle(\rho_i,\sigma_j)\rangle \subseteq S_{\mathrm{dom}(\rho_i)} \times S_{\mathrm{dom}(\sigma_j)}$ for $1 \leq i \leq r$ and $1 \leq j \leq s$. The number of orbits of $\langle(\rho_i,\sigma_j)\rangle$ is simply $gcd(|\mathrm{dom}(\rho_i)|,|\mathrm{dom}(\sigma_j)|)$ and the result follows. □

We will now define a way to construct a collection of $m \times n$ binary matrices from a partition of $\mathbf{m} \times \mathbf{n}$. This construction will play a key role for the rest of this chapter. In particular, we will eventually link various properties of $m \times n$ matrices to partitions of $\mathbf{m} \times \mathbf{n}$. This link will be such that the matrices which satisfy a property are exactly the collection of matrices constructed from the partition corresponding to that property. For example, we will later prove that the collection of matrices constructed using the relation $R_{\rho,\sigma}$ is exactly the subset of $m \times n$ binary matrices which are fixed by the pair of permutations (ρ,σ).

Without further ado, if P is a partition, or the corresponding equivalence, of $\mathbf{m} \times \mathbf{n}$ and $f : P \to \{0,1\}$ is a function we will write f' to denote the corresponding function from $\mathbf{m} \times \mathbf{n}$ to $\{0,1\}$ which sends (i,j) to the value of $[(i,j)]_P f$. When P is a partition of $\mathbf{m} \times \mathbf{n}$ and we write $\{0,1\}^P$ (normally denoting the set of all functions from P into $\{0,1\}$) we will actually mean the set $\{f' : f \in \{0,1\}^P\}$ of corresponding functions in $\{0,1\}^{\mathbf{m}\times\mathbf{n}}$. Equivalently, we can

define $\{0,1\}^P$ in terms of the kernel of a function in $\{0,1\}^{m\times n}$ by:

$$\{0,1\}^P = \{f \in \{0,1\}^{\mathbf{m}\times\mathbf{n}} : P \subseteq \ker(f)\}.$$

It is important we can represent these collections of functions as subsets of $\{0,1\}^{m\times n}$ so that we can take unions and intersections later.

We now define an equivalence relation which relates to the collection of matrices with row x equal to row y. When we refer to row r or column c of a $m\times n$ matrix f we will mean the tuples $((r,1)f,\dots(r,n)f)$ or $((1,c)f,\dots,(m,c)f)$, respectively. Let $E_{x,y}$ be the equivalence relation on $\mathbf{m}\times\mathbf{n}$ such that the non-reflexive pairs in $E_{x,y}$ are exactly those of the form $((x,z),(y,z))$ and $((y,z),(x,z))$ for all $z\in\mathbf{n}$. Then functions $f : E_{x,y}\to\{0,1\}$ correspond to $m\times n$ binary matrices with row x equal to row y. For such an f we will write f' to denote the corresponding function from $\mathbf{m}\times\mathbf{n}$ to $\{0,1\}$, that is to say $(i,j)f' = [(i,j)]_{E_{x,y}}f$ where $[(i,j)]_{E_{x,y}}$ is the class of (i,j) in $E_{x,y}$. Furthermore we define $F_{x,y}$ analogously for columns as opposed to rows.

We will want to be able to determine when matrices are in both $\{0,1\}^P$ and $\{0,1\}^Q$ for two partitions P,Q. It turns out we can describe this situation using the join $P\vee Q$ of two equivalences which is the least equivalence such that $(x,y)\in P$ and $(y,z)\in Q$ implies $(x,z)\in P\vee Q$.

Proposition 4.3.6. *Let P,Q be equivalence relations on $\mathbf{m}\times\mathbf{n}$ and write $P\vee Q$ to denote their join. Then an $m\times n$ binary matrix f satisfies $f\in\{0,1\}^{P\vee Q}$ if and only if $f\in\{0,1\}^P$ and $f\in\{0,1\}^Q$.*

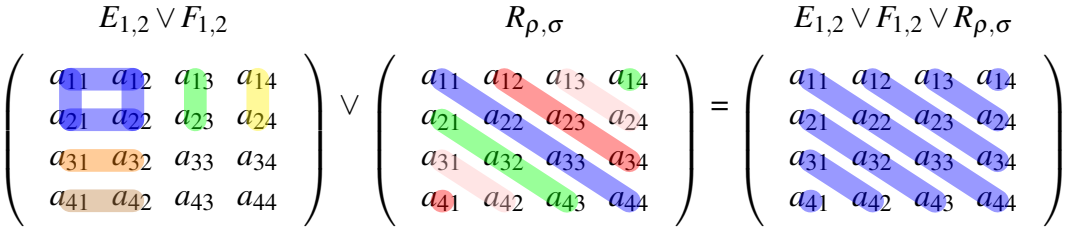
Proof. Since each equivalence class of P or Q is a subset of an equivalence class of $P\vee Q$, $\{0,1\}^{P\vee Q}$ must be a subset of both $\{0,1\}^P$ and $\{0,1\}^Q$, so lies in their intersection. On the other hand, if f lies in the intersection of $\{0,1\}^P$ and $\{0,1\}^Q$ then for every $((x_1,y_1),(x_2,y_2))$ in $P\vee Q$ we have, for some (x_3,y_3), that $((x_1,y_1),(x_3,y_3))\in\{0,1\}^P$ and $((x_3,y_3),(x_2,y_2))\in\{0,1\}^Q$. It follows that $(x_i,y_1)f = (x_3,y_3)f = (x_2,y_2)f$ and so $f\in\{0,1\}^{P\vee Q}$ as required. □

It follows that, say, to find the matrices with row a equal to row b and column c equal to column d we enumerate the collection $\{0,1\}^{E_{a,b}\vee F_{c,d}}$ and to find which of these are additionally fixed by (ρ,σ) we would enumerate the collection $\{0,1\}^{E_{a,b}\vee F_{c,d}\vee R_{\rho,\sigma}}$.

Example 4.3.7. Consider the situations just mentioned. First, $E_{1,2}\vee F_{1,2}$:

$$\underset{E_{1,2}}{\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14}\\ a_{21} & a_{22} & a_{23} & a_{24}\\ a_{31} & a_{32} & a_{33} & a_{34}\\ a_{41} & a_{42} & a_{43} & a_{44}\end{pmatrix}} \vee \underset{F_{1,2}}{\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14}\\ a_{21} & a_{22} & a_{23} & a_{24}\\ a_{31} & a_{32} & a_{33} & a_{34}\\ a_{41} & a_{42} & a_{43} & a_{44}\end{pmatrix}} = \underset{E_{1,2}\vee F_{1,2}}{\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14}\\ a_{21} & a_{22} & a_{23} & a_{24}\\ a_{31} & a_{32} & a_{33} & a_{34}\\ a_{41} & a_{42} & a_{43} & a_{44}\end{pmatrix}}$$

Second, let $\rho = \sigma = (1\,2\,3\,4)$. Then:



Again, we have highlighted entries with distinct colours indicating different orbits. Non-highlighted entries are in singleton orbits.

Using this idea we can describe those matrices which are fixed by a certain pair of permutations and have all rows and columns unique.

Corollary 4.3.8. *Let $(\rho,\sigma) \in S_m \times S_n$ and let f be an $m \times n$ binary matrix. Then f is fixed by (ρ,σ) and has all rows and columns unique if and only if*

$$f \notin \{0,1\}^{R_{\rho,\sigma} \vee E_{a,b}}$$

for any $1 \leq a < b \leq m$ and

$$f \notin \{0,1\}^{R_{\rho,\sigma} \vee F_{c,d}}$$

for any $1 \leq c < d \leq n$.

Proof. A matrix does not have all rows and columns unique precisely when there exists a,b such that row a equals row b, or there exists c,d such that column c equals column d. Therefore a matrix with all rows and columns unique is in the complement of the union of all the matrices containing a pair of equal rows or columns. □

We can enumerate $X_{m,n}^{(\rho,\sigma)}$, the set of $m \times n$ binary matrices with all rows and columns distinct fixed by (ρ,σ) by using the inclusion-exclusion principle. The inclusion-exclusion principle states that for finite sets $A_1,\ldots,A_k$ the following identity holds:

$$\left|\bigcup_{i=1}^{k} A_i\right| = \sum_{j=1}^{k} (-1)^{j+1} \left(\sum_{1 \leq i_1 < \cdots < i_j \leq k} |A_{i_1} \cap \cdots \cap A_{i_j}| \right). \tag{4.1}$$

The union of the sets

$$Y = \{\{0,1\}^{R_{\rho,\sigma} \vee E_{x,y}} : 1 \leq x < y \leq m\} \cup \{\{0,1\}^{R_{\rho,\sigma} \vee F_{x,y}} : 1 \leq x < y \leq n\}$$

is equal to the complement of $X_{m,n}^{(\rho,\sigma)}$ in $\{0,1\}^{R_{\rho,\sigma}}$. Our plan is to use the inclusion exclusion principle to determine the size of the union of the sets in Y. We can then subtract this from the size of $\{0,1\}^{R_{\rho,\sigma}}$, which we can calculate, to obtain the size of $X_{m,n}^{(\rho,\sigma)}$.

However the set Y is larger than it needs to be. Some sets in Y are subsets of other sets in Y so removing them will not change the union, and it will simplify the calculation involved in applying the inclusion-exclusion principle. This corresponds to discounting the equivalences which are finer than others. If $P \supseteq Q$ then $\{0,1\}^P$ contains $\{0,1\}^Q$ and therefore we will discard $\{0,1\}^Q$ from Y. We next observe one of the primary situations where this happens. However we will first define some convenient notation. Herein we will write $\rho_1,\ldots,\rho_r$ to denote the distinct cycles of the permutation called $\rho \in S_m$, including those of length one for elements in $\mathbf{m}$ which ρ fixes. Similarly, we will write $\sigma_1,\ldots,\sigma_s$ to denote the distinct cycles of the permutation called $\sigma \in S_n$, again including those of length one. Also note that we denote the number of disjoint cycles of ρ and σ by r and s, respectively. We will consider the domain of a cycle to be just those elements in the cycle it refers to, i.e. the domain will be of size equal to the length of that cycle rather than m or n.

Lemma 4.3.9. *Let $(\rho,\sigma) \in S_m \times S_n$. Let ρ_i be a cycle of ρ. Let $x,y \in \mathrm{dom}(\rho_i)$. Let $a,b \in |\rho_\mathbf{i}|$. Then $E_{x,x\rho^a} \vee R_{\rho,\sigma}$ is contained in $E_{y,y\rho^b} \vee R_{\rho,\sigma}$ if $\gcd(|\rho_i|,a)$ divides $\gcd(|\rho_i|,b)$. An analogous result holds for the corresponding equivalences for columns.*

Proof. Let $g_a = \gcd(|\rho_i|,a)$ divide $g_b = \gcd(|\rho_i|,b)$. The result follows from showing $E_{x,x\rho^a} \vee R_{\rho,\sigma}$ is coarser than $E_{y,y\rho^b}$ which is equivalent to showing $((y,z),(y\rho^b,z)$ is in $E_{x,x\rho^a} \vee R_{\rho,\sigma}$ for all $z \in \mathbf{n}$. Let $z \in \mathbf{n}$ and choose k such that $x\rho^k = y$ and let $w = z^{-k}$. Since g_a divides g_b there exists λ such that $b = \lambda a \mod |\rho_i|$. Therefore we have

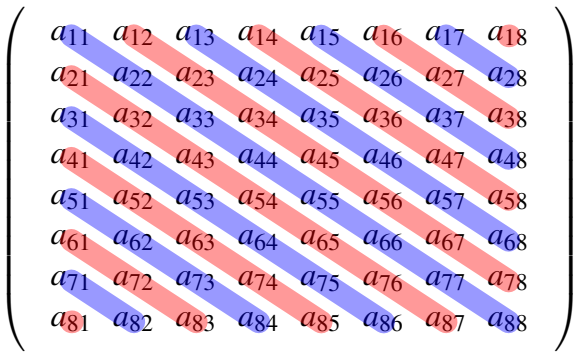
$$
\begin{aligned}
(y,z)R_{\rho,\sigma}(y\rho^{-k}, z\sigma^{-k}) = (x,w)\\
&E_{x,x\rho^a}(x\rho^a, w)\\
&R_{\rho,\sigma}(x, w\sigma^{-a})\\
&E_{x,x\rho^a}(x\rho^a, w\sigma^{-a})\\
&R_{\rho,\sigma}(x\rho^{2a}, w)\\
&\vdots\\
&R_{\rho,\sigma}(x\rho^{3a}, w)\\
&\vdots\\
&R_{\rho,\sigma}(x\rho^{\lambda a}, w) = (x\rho^b, w)\\
&R_{\rho,\sigma}(x\rho^{b+k}, w\sigma^k) = (y\rho^b, z)
\end{aligned}
$$

which holds for all z so we are done. The analogous result for column equivalences follows by a dual argument. □

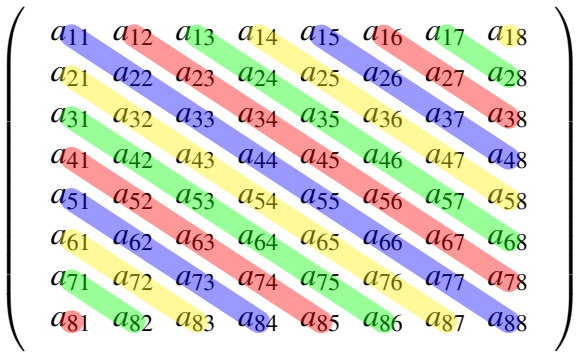
Therefore we will write $E_{\rho_i,k}$ to denote the equivalence which is equal to $E_{x,x\rho^k} \vee R_{\rho,\sigma}$, which is the same equivalence for any choice of $x \in \mathrm{dom}(\rho_i)$. The unique $E_{\rho_i,k}$ are in correspondence with the divisors of $|\rho_i|$ and we will herein only write $E_{\rho_i,k}$ when k is a divisor of $|\rho_i|$. We note that the matrices in the set $\{0,1\}^{E_{\rho_i,k}}$ can equivalently be described as those fixed by (ρ,σ) such that row x equals row x^k for all $x \in \mathrm{dom}(\rho_i)$. We define equivalences of the form $F_{\sigma_j,k}$ analogously.

Example 4.3.10. Let $\rho = \sigma = (1\,2\,3\,4\,5\,6\,7\,8) \in S_8$. Then Lemma 4.3.9 tells us that, say, $E_{1,3} \vee R_{\rho,\sigma}$ is contained in, say, $E_{1,5} \vee R_{\rho,\sigma}$ since $3 = 1\rho^2$, $5 = 1\rho^4$, and $\gcd(2,8)$ divides $\gcd(4,8)$. We now prefer to refer to these equivalences as $E_{\rho,2}$ and $E_{\rho,4}$, respectively. We picture them as those matrices fixed by (ρ,σ) such that row x equals rows x^{ρ^k} modulo 8 for $k = 2,4$ (respectively). In this case row x equals row $x+k$ modulo 8. Note we consider row 0 to be represent row 8.

$E_{\rho,2}$



$E_{\rho,4}$



We will next show that an equivalence of the form $E_{x,y} \vee R_{\rho,\sigma}$ where $x \in \mathrm{dom}(\rho_i)$, $y \in \mathrm{dom}(\rho_j)$ and $|\rho_i| \neq |\rho_j|$ must be coarser than some equivalence of the form $E_{\rho_z,k}$. Thus we will be able to discount them as discussed earlier.

Lemma 4.3.11. *Let $(\rho,\sigma) \in S_m \times S_n$. Let ρ_i, ρ_j be cycles of ρ such that $|\rho_i| \neq |\rho_j|$. Let $k = gcd(|\rho_i|,|\rho_j|)$. Let $x \in \mathrm{dom}(\rho_i)$ and $y \in \mathrm{dom}(\rho_j)$. If $k < |\rho_i|$ then $E_{x,y} \vee R_{\rho,\sigma}$ is coarser than $E_{x,x\rho^k} \vee R_{\rho,\sigma}$; and if $k < |\rho_j|$ then $E_{x,y} \vee R_{\rho,\sigma}$ is coarser than $E_{y,y\rho^k} \vee R_{\rho,\sigma}$. An analogous result holds for the corresponding equivalences for columns.*

Proof. Let $k = gcd(|\rho_i|,|\rho_j|)$ and, without loss of generality, assume $k < |\rho_i|$ holds. Then we aim to show that $E_{x,y} \vee R_{\rho,\sigma}$ is coarser than $E_{x,x\rho^k}$ by showing that $(x,z)E_{x,y} \vee R_{\rho,\sigma}(x\rho^k,z)$ for all $z \in \mathbf{n}$. Let $z \in \mathbf{n}$ and suppose that we may write $k = \alpha|\rho_i| + \beta|\rho_j|$ (which we can do using

the euclidean algorithm) then

$$
\begin{aligned}
&(x,z)R_{\rho,\sigma}(x\rho^{-|\rho_i|}, z\sigma^{-|\rho_i|}) = (x, z\sigma^{-|\rho_i|})\\
&\quad R_{\rho,\sigma}(x, z\sigma^{-2|\rho_i|})\\
&\quad \dots\\
&\quad R_{\rho,\sigma}(x, z\sigma^{-\alpha|\rho_i|})\\
&\quad E_{x,y}(y, z\sigma^{-\alpha|\rho_i|})\\
&\quad R_{\rho,\sigma}(y\rho^{-|\rho_j|}, z\sigma^{-\alpha|\rho_i|-|\rho_j|}) = (y, z\sigma^{-\alpha|\rho_i|-|\rho_j|})\\
&\quad R_{\rho,\sigma}(y, z\sigma^{-\alpha|\rho_i|-2|\rho_j|})\\
&\quad \dots\\
&\quad R_{\rho,\sigma}(y, z\sigma^{-\alpha|\rho_i|-\beta|\rho_j|}) = (y, z\sigma^{-k})\\
&\quad E_{x,y}(x, z\sigma^{-k})\\
&\quad R_{\rho,\sigma}(z\rho^{k}, z)
\end{aligned}
$$

as required. Since this holds for all z then $E_{x,y} \vee R_{\rho,\sigma}$ must be coarser than $E_{x,x\rho^k}$. It follows immediately that $E_{x,y} \vee R_{\rho,\sigma}$ must be coarser than $E_{x,x\rho^k} \vee R_{\rho,\sigma}$. If $k < |p_j|$ then we may show that $E_{x,y} \vee R_{\rho,\sigma}$ is coarser than $E_{y,y\rho^k} \vee R_{\rho,\sigma}$ by a similar argument. Finally, the result for column equivalences follows an argument analogous to that just presented. □

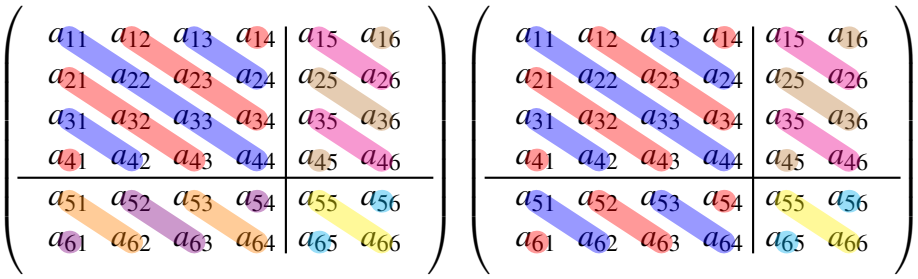
We have now narrowed down the equivalences we must consider to those of the form $E_{\rho_i,k}$ and those of the form $E_{x,y} \vee R_{\rho,\sigma}$ for $x \in \mathrm{dom}(\rho_i)$, $y \in \mathrm{dom}(\rho_j)$ where $i \neq j$ and $|\rho_i| = |\rho_j|$.

Example 4.3.12. Let $\rho = \sigma = (1\,2\,3\,4)(5\,6) \in S_6$. Then Lemma 4.3.11 tells us that $E_{1,5} \vee R_{\rho,\sigma}$ is coarser than $E_{1,3} \vee R_{\rho,\sigma}$.

$R_{\rho,\sigma}$

$$
\left(\begin{array}{cccc|cc}
a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16}\\
a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26}\\
a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36}\\
a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46}\\
\hline
a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56}\\
a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66}
\end{array}\right)
$$

$E_{1,3} \vee R_{\rho,\sigma}$ $E_{1,5} \vee R_{\rho,\sigma}$



For equivalences of the form $E_{x,y} \vee R_{\rho,\sigma}$ we will find it easier to consider several of these cases simultaneously. We will define

$$\{0,1\}^{E_{\rho_i,\rho_j}} := \bigcup_{x \in \mathrm{dom}(\rho_i)} \bigcup_{y \in \mathrm{dom}(\rho_j)} \{0,1\}^{E_{x,y} \vee R_{\rho,\sigma}}$$

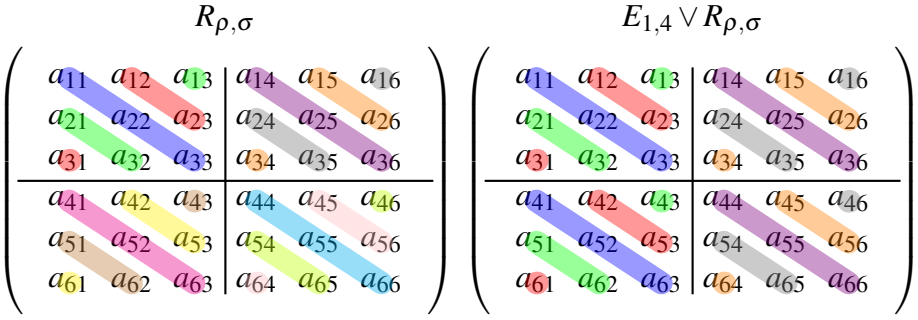
which is a union of the sets of matrices corresponding to multiple different equivalences linking the cycles ρ_1 and ρ_2. In a similar manner, we define:

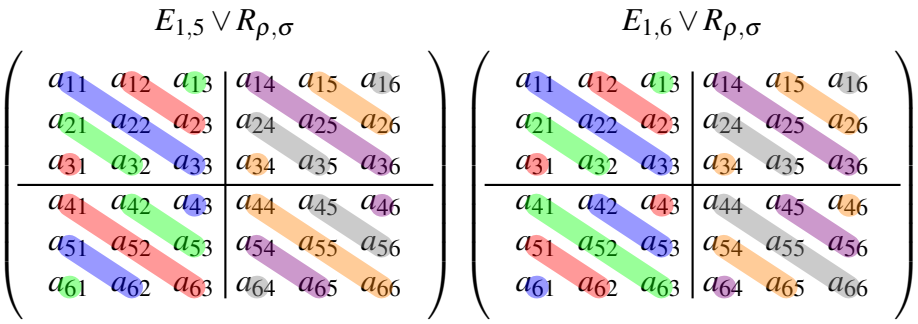
$$\{0,1\}^{F_{\sigma_i,\sigma_j}} := \bigcup_{x \in \mathrm{dom}(\sigma_i)} \bigcup_{y \in \mathrm{dom}(\sigma_j)} \{0,1\}^{F_{x,y} \vee R_{\rho,\sigma}}.$$

We will find these collections to be convenient when it comes to applying the inclusion-exclusion principle, as opposed to using all of the sets $\{0,1\}^{E_{x,y} \vee R_{\rho,\sigma}}$. Note however that there may not be an equivalence relation corresponding to this union, so the continued use of the power notation $\{0,1\}^{E_{\rho_i,\rho_j}}$ is misleading. We can think of the matrices in this set as the subset of those matrices fixed by (ρ,σ) such that the block of rows corresponding to ρ_i and ρ_j are equal up to a permutation which preserves the cycle structure.

Example 4.3.13. Let $\rho = \sigma = (123)(456) \in S_6$ and denote $\rho_1 = (123)$ and $\rho_2 = (456)$. Then E_{ρ_1,ρ_2} is equal to the union of the sets $\{0,1\}^{E_{1,4} \vee R_{\rho,\sigma}}$, $\{0,1\}^{E_{1,5} \vee R_{\rho,\sigma}}$, and $\{0,1\}^{E_{1,6} \vee R_{\rho,\sigma}}$.

$R_{\rho,\sigma}$ $E_{1,4} \vee R_{\rho,\sigma}$





For example if $\rho_i = (1,2,3)$ and $\rho_j = (4,5,6)$ then a matrix in this set will satisfy row 1 equals row z for some $z \in \{4,5,6\}$. Whichever z is chosen, it must also be true that row $1\rho_i$ equals row $z\rho_j$ and row $1\rho_i^2$ equals row $z\rho_j^2$. We now present this idea in generality.

Lemma 4.3.14. *Let $(\rho,\sigma) \in S_m \times S_n$. Let $f \in \{0,1\}^{R_{\rho,\sigma}}$. Let ρ_i, ρ_j be distinct cycles of ρ such that $|\rho_i| = |\rho_j|$. Then $f \in \{0,1\}^{E_{\rho_i,\rho_j}}$ if and only if there is a map $\theta : \mathrm{dom}(\rho_i) \to \mathrm{dom}(\rho_j)$ such that for all $z \in \mathrm{dom}(\rho_i)$ the following two statements are true:*

(i) *row z of f and row $\theta(z)$ of f are equal, and*

(ii) *$\theta(z\rho)$ is equal to $(\theta(z))\rho$.*

The analogous result in terms of columns is also true.

Proof. Assume $f \in \{0,1\}^{E_{\rho_i,\rho_j}}$ and we will prove the forward implication. By assumption, there exists $x \in \mathrm{dom}(\rho_i)$ and $y \in \mathrm{dom}(\rho_j)$ such that $f \in \{0,1\}^{E_{x,y} \vee R_{\rho,\sigma}}$. Therefore row x of f is equal to row y of f. We will choose $\theta : \mathrm{dom}(\rho_i) \to \mathrm{dom}(\rho_j)$ to be the map such that $\theta(x) = y$ and $\theta(x\rho^k) = y\rho^k$ for all $k \in \mathbb{Z}$. This is well defined since $|\rho_i|$ equals $|\rho_j|$. By definition θ satisfies (ii) and we will now show it also satisfies (i). Consider $((x,1)f,\ldots,(x,n)f)$ which is row x of f. Let $k \in \mathbb{Z}$. Then since f is fixed by (ρ,σ) we have that row $x\rho^k$ is equal to $((x,1\sigma^{-k})f,\ldots,(x,n\sigma^{-k})f)$. Furthermore row y is equal to row x and it follows that row $y\rho^k$ is equal to $((x,1\sigma^{-k})f,\ldots,(x,n\sigma^{-k})f)$ which is equal to row $x\rho^k$. Since every $z \in \mathrm{dom}(\rho_i)$ can be expressed in the form $x\rho^k$ for some $k \in \mathbb{Z}$ we have shown (i).

For the reverse implication we assume that there exists a map θ as described in the lemma statement. Let $z \in \mathrm{dom}(\rho_i)$ then row z is equal to row $\theta(z)$ by assumption. Therefore $f \in \{0,1\}^{E_{x,y}}$. We also know that $f \in \{0,1\}^{R_{\rho,\sigma}}$. Recall that

$$\{0,1\}^{E_{\rho_i,\rho_j}} := \bigcup_{x \in \rho_i} \bigcup_{y \in \rho_j} \{0,1\}^{E_{x,y} \vee R_{\rho,\sigma}}.$$

Thus we deduce

$$f \in \{0,1\}^{E_{x,y}} \cap \{0,1\}^{R_{\rho,\sigma}} = \{0,1\}^{E_{x,y} \vee R_{\rho,\sigma}} \subseteq \{0,1\}^{E_{\rho_i,\rho_j}}$$

(the equality holds due to Proposition 4.3.6). This completes the proof. The result for columns is shown by an analogous argument. □

The function θ in Lemma 4.3.14 essentially describes which set of the form $\{0,1\}^{E_{x,y}\vee R_{\rho,\sigma}}$ a matrix $f \in \{0,1\}^{E_{\rho_i,\rho_j}}$ belongs to. To be more specific, if $x \in \mathrm{dom}(\rho_i)$ and $y \in \mathrm{dom}(\rho_j)$ then $y = \theta(x)$ implies $f \in \{0,1\}^{E_{x,y}\vee R_{\rho,\sigma}}$.

4.3.1 Properties of sub-matrices

Let f be a $m \times n$ matrix. Then if $I \subseteq \mathbf{m}$ and $J \subseteq \mathbf{n}$ we can define a sub-matrix of f on the domain $I \times J$. We will write $f|_{I\times J}$ to denote the restriction of f to the domain $I \times J$ and we will call this restriction a *sub-matrix* of f. Let $(\rho,\sigma) \in S_m \times S_n$ and assume $f \in \{0,1\}^{R_{\rho,\sigma}}$. We will primarily be concerned with sub-matrices where the domain $I \times J$ is invariant under the natural action of (ρ,σ). That is to say $\{(x\rho, y\sigma) : (x,y) \in I \times J\}$ is equal to $I \times J$. Next we define a property of sub-matrices of matrices fixed by a pair of permutations which will be key to our enumeration strategy in the subsequent section.

Let $(\rho,\sigma) \in S_m \times S_n$. Let f be a $m \times n$ matrix. Let $I \subseteq \mathbf{m}$ and $J \subseteq \mathbf{n}$ such that $I \times J$ is invariant under the action of (ρ,σ). Let $p \in \mathbb{N}$ be a divisor of $|I|$. Then we say that the sub-matrix $f|_{I\times J}$ of f has *row period* p with respect to (ρ,σ) if for all $(x,y) \in I \times J$ we have that $(x,y)f = (x\rho^z, y)f$ if and only if there exists $k \in \mathbb{Z}$ such that $z = kp$. We define *column period* analogously.

Herein we will neglect to write 'with respect to (ρ,σ)' when referring to the row period or column period of a sub-matrix. The respective pair of permutations will always be named (ρ,σ) and will be clear in context. We will exclusively be interested in the row period of cases where $I = \mathrm{dom}(\rho_i)$ is the domain of a cycle of ρ. Similarly, we will only care about the column period of cases where $J = \mathrm{dom}(\sigma_j)$ is the domain of a cycle of σ. Recalling the figure from Example 4.3.4 will be helpful to visualise these sub-matrices.

Example 4.3.15. The following examples where $\rho = \sigma = (1\,2\,3\,4)$ show matrices with row period 'RP' and column period 'CP' equal to 1, 2, and 4. We can see they are also fixed by (ρ,σ) because entries highlighted by the same colour (that is, those that have indices related by $R_{\rho,\sigma}$) are all equal.

RP = CP = 1 RP = CP = 2 RP = CP = 4

$$\begin{pmatrix} 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{pmatrix} \begin{pmatrix} 1&0&1&0\\ 0&1&0&1\\ 1&0&1&0\\ 0&1&0&1 \end{pmatrix} \begin{pmatrix} 0&0&0&1\\ 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0 \end{pmatrix}.$$

	$f\|_{\mathrm{dom}(\rho_1)\times\mathrm{dom}(\sigma_1)}$	$f\|_{\mathrm{dom}(\rho_1)\times\mathrm{dom}(\sigma_2)}$	$f\|_{\mathrm{dom}(\rho_2)\times\mathrm{dom}(\sigma_1)}$	$f\|_{\mathrm{dom}(\rho_2)\times\mathrm{dom}(\sigma_2)}$
RP	4	1	2	1
CP	4	1	2	1

	$f\|_{\mathrm{dom}(\rho_1)\times\mathbf{n}}$	$f\|_{\mathrm{dom}(\rho_2)\times\mathbf{n}}$	$f\|_{\mathbf{m}\times\mathrm{dom}(\sigma_1)}$	$f\|_{\mathbf{m}\times\mathrm{dom}(\sigma_2)}$
RP	4	2	n/a	n/a
CP	n/a	n/a	4	1

More interesting examples occur when both ρ and σ are not cycles. Let $\rho = (1234)(56)$ and $\sigma = (1234)(5678)$. Then consider the following matrix which is in $\{0,1\}^{R_{\rho,\sigma}}$. We have tabulated the row and column periods of eight sub-matrices.

$$\begin{array}{c} \\ \rho_1 \\ \\ \rho_2 \\ \\ \end{array} \begin{array}{c} \begin{array}{cccc} \sigma_1 & & & \end{array} \quad \begin{array}{cccc} \sigma_2 & & & \end{array} \\ \left(\begin{array}{cccc|cccc} 0&0&0&1&1&1&1&1\\ 1&0&0&0&1&1&1&1\\ 0&1&0&0&1&1&1&1\\ 0&0&1&0&1&1&1&1\\ \hline 1&0&1&0&0&0&0&0\\ 0&1&0&1&0&0&0&0 \end{array}\right) \end{array}.$$

In the case of a sub-matrix on a domain $\mathrm{dom}(\rho_i) \times \mathrm{dom}(\sigma_j)$ we can show that row period and column period refer to the same property.

Lemma 4.3.16. *Let $(\rho,\sigma) \in S_m \times S_n$. Let $f \in \{0,1\}^{R_{\rho,\sigma}}$. Let ρ_i be a cycle of ρ and let σ_j be a cycle of σ. Then the row period and the column period of $f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)}$ are equal.*

Proof. Let $(x,y) \in \mathrm{dom}(\rho_i) \times \mathrm{dom}(\sigma_j)$. Assume that the row period of $f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)}$ is equal to p. Then $(x\rho^{-p},y)f = (x,y)$. Since $f \in \{0,1\}^{R_{\rho,\sigma}}$ it also true that $(x\rho^{-p},y)f = (x\rho^{-p}\rho^p, y\sigma^p)f$. Thus we deduce that

$$(x,y)f = (x\rho^{-p},y)f = (x\rho^{-p}\rho^p, y\sigma^p) = (x,y\sigma^p)f.$$

Since (x,y) was arbitrary in $\mathrm{dom}(\rho_i) \times \mathrm{dom}(\sigma_j)$, we have that p divides the column period of $f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)}$. An analogous argument shows that the column period of $f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)}$ divides the row period of $f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)}$, and so they must be equal. □

The result of Lemma 4.3.16 is demonstrated in Example 4.3.15. Lemma 4.3.16 will also play a part in proving the next result. Our next Lemma shows how certain sub-matrices with domains equal to the product of a cycle of ρ with a cycle of σ must have equal row periods.

Lemma 4.3.17. *Let $(\rho,\sigma) \in S_m \times S_n$. Let ρ_i, ρ_j be distinct cycles of ρ such that $|\rho_i| = |\rho_j|$. Let $f \in \{0,1\}^{E_{\rho_i,\rho_j}}$. Then for any cycle σ_k of σ the sub-matrices $f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_k)}$ and $f|_{\mathrm{dom}(\rho_j)\times\mathrm{dom}(\sigma_k)}$ have equal row period. Similarly, let σ_i, σ_j be distinct cycles of σ. If $f \in \{0,1\}^{F_{\sigma_i,\sigma_j}}$ then for any cycle ρ_k of ρ the sub-matrices $f|_{\mathrm{dom}(\rho_k)\times\mathrm{dom}(\sigma_i)}$ and $f|_{\mathrm{dom}(\rho_k)\times\mathrm{dom}(\sigma_j)}$ have equal column period.*

Proof. First, let $f \in \{0,1\}^{E_{\rho_i,\rho_j}}$. Assume that the row period of $f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_k)}$ is equal to p and that the row period of $f|_{\mathrm{dom}(\rho_j)\times\mathrm{dom}(\sigma_k)}$ is equal to q. Lemma 4.3.14 shows $f \in \{0,1\}^{E_{\rho_i,\rho_j}}$ implies there is a map $\theta : \mathrm{dom}(\rho_i) \to \mathrm{dom}(\rho_j)$ such that for all $z \in \mathrm{dom}(\rho_i)$ we have $\theta(z\rho)$ is equal to $(\theta(z))\rho$, and row z is equal to row $\theta(z)$. Let $y \in \mathrm{dom}(\rho_j)$ and denote $x = \theta^{-1}(y)$. Then row x is equal to row $x\rho^p$ by the definition of θ. Furthermore, row x is equal to row $y = \theta(x)$ and row $x\rho^p$ is equal to row $\theta(x\rho^p)$. Moreover, by definition of θ we have that $\theta(x\rho^p)$ is equal to $(\theta(x))\rho^p = y\rho^p$. Therefore row $x\rho^p$ is equal to row $y\rho^p$. Thus we have shown row y is equal to row x which was assumed equal to row $x\rho^p$ which is equal to row $y\rho^p$. Since y was arbitrary in $\mathrm{dom}(\rho_j)$ we have that p divides the row period q of $f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_k)}$. A dual argument shows that q divides p. Thus the two row periods are equal. The result for column periods follows from an analogous argument. □

We now state a corollary of Lemma 4.3.17 which gives a more general statement about which of these types of sub-matrices have equal row period.

Corollary 4.3.18. *Let $(\rho,\sigma) \in S_m \times S_n$. Let ρ_i, ρ_j be distinct cycles of ρ such that $|\rho_i| = |\rho_j|$. Let $f \in \{0,1\}^{E_{\rho_i,\rho_j} \vee F_{\sigma_k,\sigma_l}}$. Then the sub-matrices $f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_k)}$ and $f|_{\mathrm{dom}(\rho_j)\times\mathrm{dom}(\sigma_l)}$ have equal row period.*

Proof. Lemma 4.3.17 tells us that $f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_k)}$ has the same row period as $f|_{\mathrm{dom}(\rho_j)\times\mathrm{dom}(\sigma_k)}$, and that $f|_{\mathrm{dom}(\rho_j)\times\mathrm{dom}(\sigma_k)}$ has the same column period as $f|_{\mathrm{dom}(\rho_j)\times\mathrm{dom}(\sigma_l)}$. Lemma 4.3.16 tells us that the row period and column period of $f|_{\mathrm{dom}(\rho_j)\times\mathrm{dom}(\sigma_k)}$ are equal. The result follows. □

Lemma 4.3.17 also allows us to quickly prove the following result about the row periods of sub-matrices whose domain spans all n columns of the whole matrix.

Corollary 4.3.19. *Let $(\rho,\sigma) \in S_m \times S_n$. Let ρ_i, ρ_j be distinct cycles of ρ such that $|\rho_i| = |\rho_j|$. Let $f \in \{0,1\}^{E_{\rho_i,\rho_j}}$. Then the sub-matrices $f|_{\mathrm{dom}(\rho_i)\times\mathbf{n}}$ and $f|_{\mathrm{dom}(\rho_j)\times\mathbf{n}}$ have equal row period. An analogous result holds for columns instead of rows.*

Proof. The row period of $f|_{\mathrm{dom}(\rho_i)\times\mathbf{n}}$ is equal to the least common multiple of the row periods of $\{f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_k)} : k \in \mathbf{s}\}$, where s is the number of cycles of σ. Lemma 4.3.17 tells us that $f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_k)}$ has the same row period as $f|_{\mathrm{dom}(\rho_j)\times\mathrm{dom}(\sigma_k)}$ for all cycles σ_k of σ. The result follows immediately. □

Finally we take a moment to deduce what Lemma 4.3.9 tells us about matrices in some $\{0,1\}^{E_{\rho_i,k}}$ in terms of row periods of certain sub-matrices.

Lemma 4.3.20. *Let $(\rho,\sigma) \in S_m \times S_n$. Let $f \in \{0,1\}^{R_{\rho,\sigma}}$. Let ρ_i be a cycle of ρ. Let k divide $|\rho_i|$. Then $f \in \{0,1\}^{E_{\rho_i,k}}$ if and only if the sub-matrix $f|_{\mathrm{dom}(\rho_i)\times\mathbf{n}}$ has row period dividing k. An analogous result holds for the equivalences relating to columns.*

Proof. Assume $f \in \{0,1\}^{E_{\rho_i,k}}$ and let $x \in \mathrm{dom}(\rho_i)$. Then $E_{x,x(\rho^k)} \vee R_{\rho,\sigma}$ is coarser than $E_{\rho_i,k}$ by Lemma 4.3.9. Therefore f is in $\{0,1\}^{E_{x,x(\rho^k)}\vee R_{\rho,\sigma}}$ which is a subset of $\{0,1\}^{E_{x,x(\rho^k)}}$ by Proposition 4.3.6. The fact that f is in $\{0,1\}^{E_{x,x(\rho^k)}}$ implies row x equals row x^k. Since x was arbitrary in $\mathrm{dom}(\rho_i)$, the row period of $f|_{\mathrm{dom}(\rho_i)\times\mathbf{n}}$ must divide k.

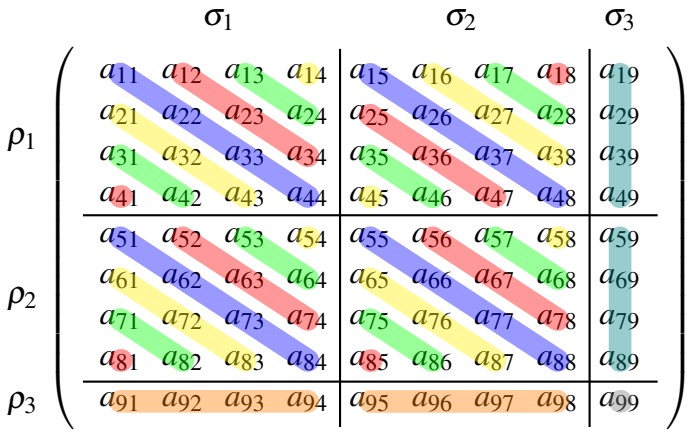
To prove the reverse implication, assume row x of f equals row x^k for all $x \in \mathrm{dom}(\rho_i)$. Then certainly $f \in \{0,1\}^{E_{x,x(\rho^k)}}$. By assumption $f \in \{0,1\}^{R_{\rho,\sigma}}$. Therefore

$$f \in \{0,1\}^{R_{\rho,\sigma}} \cap \{0,1\}^{E_{x,x(\rho^k)}} = \{0,1\}^{R_{\rho,\sigma}\vee E_{x,x(\rho^k)}} = \{0,1\}^{E_{\rho_i,k}}$$

as required. The analogous result for columns follows by an identical argument. □

We have now sufficiently developed our theory of equivalences to proceed to the enumeration strategy. We end this section with an example to illustrate the last few results.

Example 4.3.21. Let $\rho = \sigma = (1234)(5678)(9) \in S_9$. Denote the cycles of ρ,σ by $\rho_1 = \sigma_1 = (1234)$, $\rho_2 = \sigma_2 = (5678)$, and $\rho_3 = \sigma_3 = (9)$. The following figure illustrates the equivalence classes of $E_{1,5} \vee F_{1,5} \vee R_{\rho,\sigma}$. A matrix in $\{0,1\}^{E_{1,5}\vee F_{1,5}\vee R_{\rho,\sigma}}$ arises from assigning 0 or 1 to each class (each colour in the illustration). We can see that for such a matrix the row period of the sub-matrix with domain $\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)$ is the same for all $i,j \in \{1,2\}$ which agrees with the results of Lemma 4.3.17 and Corollary 4.3.18. Furthermore the row period of the sub-matrices with domains $\mathrm{dom}(\rho_1)\times\mathbf{n}$ and $\mathrm{dom}(\rho_2)\times\mathbf{n}$ are equal, confirming Corollary 4.3.19.


σ1
σ2
σ3
ρ1
ρ2
ρ3
a11 a12 a13 a14 a15 a16 a17 a18 a19
a21 a22 a23 a24 a25 a26 a27 a28 a29
a31 a32 a33 a34 a35 a36 a37 a38 a39
a41 a42 a43 a44 a45 a46 a47 a48 a49
a51 a52 a53 a54 a55 a56 a57 a58 a59
a61 a62 a63 a64 a65 a66 a67 a68 a69
a71 a72 a73 a74 a75 a76 a77 a78 a79
a81 a82 a83 a84 a85 a86 a87 a88 a89
a91 a92 a93 a94 a95 a96 a97 a98 a99

4.4 Enumeration

To summarise the previous section, we have defined four important constructions of sets of matrices for each pair of permutations $(\rho,\sigma) \in S_m \times S_n$:

(i) for each ρ_i a cycle of ρ and k a proper divisor of $|\rho_i|$ we have $\{0,1\}^{E_{\rho_i,k}}$ where $E_{\rho_i,k}$ is defined to be equal to $E_{x,x\rho^k} \vee R_{\rho,\sigma}$ where x is any element of $\mathrm{dom}(\rho_i)$;

(ii) for each pair ρ_i,ρ_j of distinct cycles of ρ such that $|\rho_i| = |\rho_j|$ we have $\{0,1\}^{E_{\rho_i,\rho_j}}$ which is defined to be equal to

$$\bigcup_{(x,y)\in \mathrm{dom}(\rho_i)\times \mathrm{dom}(\rho_j)} \{0,1\}^{E_{x,y}\vee R_{\rho,\sigma}};$$

(iii) for each σ_i a cycle of σ and k a proper divisor of $|\sigma_i|$ we have $\{0,1\}^{F_{\sigma_i,k}}$ where $F_{\sigma_i,k}$ is defined to be equal to $F_{x,x\sigma^k} \vee R_{\rho,\sigma}$ where x is any element of $\mathrm{dom}(\sigma_i)$;

(iv) for each pair σ_i,σ_j of distinct cycles of σ such that $|\sigma_i| = |\sigma_j|$ we have $\{0,1\}^{F_{\sigma_i,\sigma_j}}$ which is defined to be equal to

$$\bigcup_{(x,y)\in \mathrm{dom}(\sigma_i)\times \mathrm{dom}(\sigma_j)} \{0,1\}^{F_{x,y}\vee R_{\rho,\sigma}}.$$

In fact if we consider the sets $(i),(ii)$ for the pair $(\sigma,\rho) \in S_n \times S_m$ and transpose the matrices in these sets then we have the sets of type $(iii),(iv)$ from the case $(\rho,\sigma) \in S_m \times S_n$. Similarly, if we consider the sets $(iii),(iv)$ for the pair (σ,ρ) and transpose the matrices in these sets then we have the sets of type $(i),(ii)$ from the case (ρ,σ). Therefore when we determine properties of $(i),(ii)$ we will not duplicate our effort by proving how to determine the equivalent properties of $(iii),(iv)$. Recall that we write $\rho_1,\ldots,\rho_r$ to denote the distinct cycles of ρ and $\sigma_1,\ldots,\sigma_s$ to denote the distinct cycles of σ. This notation includes cycles of length one for the fixed points of ρ and σ.

We define the function $\mathscr{A}$ which maps any pair of permutations $(\rho,\sigma) \in S_m \times S_n$ to the set containing all sets of matrices of types $(i)-(iv)$ which is denoted $\mathscr{A}(\rho,\sigma)$. Our aim is to apply the inclusion-exclusion principle to the collection $\mathscr{A}(\rho,\sigma) = \{a_1,a_2,\ldots,a_{|\mathscr{A}(\rho,\sigma)|}\}$. The inclusion-exclusion principle tells us that:

$$\left|\bigcup_{a_i\in\mathscr{A}(\rho,\sigma)} a_i\right| = \sum_{\substack{1\le k\le|\mathscr{A}(\rho,\sigma)|\\ 1\le i_1<i_2<\ldots<i_k\le|\mathscr{A}(\rho,\sigma)|}} (-1)^{k+1}|a_{i_1}\cap a_{i_2}\cap\cdots\cap a_{i_k}|. \tag{4.2}$$

As discussed after Corollary 4.3.8, this union is the complement of $X_{m,n}^{\rho,\sigma}$ in $\{0,1\}^{R_{\rho,\sigma}}$. The size of $\{0,1\}^{R_{\rho,\sigma}}$ is known and so we can use Equation 4.2 to deduce $|X_{m,n}^{\rho,\sigma}|$. We can simplify the expression in Equation 4.2 to something easier to compute. The majority of the computational effort will be spent determine the sizes of the intersections of subsets of $\mathscr{A}(\rho,\sigma)$. Typically there are many subsets of $\mathscr{A}(\rho,\sigma)$ which have intersections of equal size and it will simplify our calculations if we can restate the expression on the right hand side of Equation 4.2 by collecting terms which have intersections of equal size. To demonstrate this simplification in generality, let P be some partition of the power set $\mathscr{P}(\mathscr{A}(\rho,\sigma))$ of $\mathscr{A}(\rho,\sigma)$ where for any $p \in P$ if $A, B \in p$, then

$$\left|\bigcap A\right| = \left|\bigcap B\right|.$$

We then define the coefficients k_p for all $p \in P$ by

$$k_p = \sum_{A \in p} (-1)^{|A|+1}. \tag{4.3}$$

These coefficients are the sum over elements of the part p of P where we add 1 for an element of the power set of $\mathscr{A}(\rho,\sigma)$ with odd cardinality and -1 for an element with even cardinality. For every p in P let us specify some element A_p of p. Then, without loss of generality, we can restate the expression given by the inclusion-exclusion principle in Equation 4.2 as:

$$\left|\bigcup_{a \in \mathscr{A}(\rho,\sigma)} a\right| = \sum_{p \in P} k_p \left|\bigcap A_p\right|. \tag{4.4}$$

We will now create a representation for elements of $\mathscr{P}(\mathscr{A}(\rho,\sigma))$ which will lead to a partition which we can use in the way just described.

4.4.1 A graphical representation

In this section we invent a representation of $\mathscr{A}(\rho,\sigma)$ involving labelled graphs which will prove useful in our enumeration. We will also define properties of this representation and prove some results which will be important in later sections.

Definition 4.4.1. Let $(\rho,\sigma) \in S_m \times S_n$. If A is a subset of $\mathscr{A}(\rho,\sigma)$ then we define the corresponding pair of graphs with labelled vertices

$$G(A) = (G_R(A), G_C(A))$$

as follows:

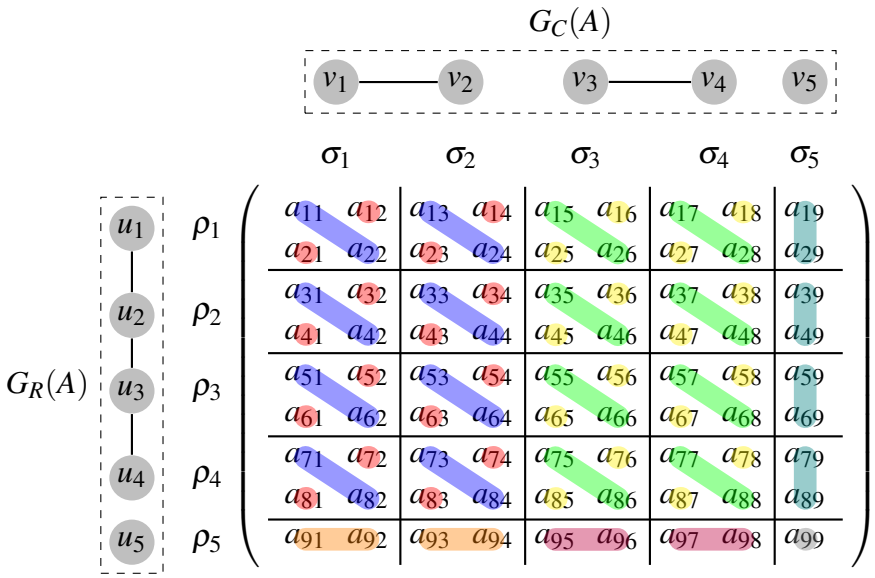


Fig. 4.1 The graph pair corresponding to $\{\{0,1\}^{E_{\rho_1,\rho_2}}, \{0,1\}^{E_{\rho_2,\rho_3}}, \{0,1\}^{E_{\rho_3,\rho_4}}, \{0,1\}^{F_{\sigma_1,\sigma_2}}, \{0,1\}^{F_{\sigma_1,\sigma_2}}\}$.

(i) The graph $G_R(A)$ has r vertices where r is the number of cycles of ρ. The vertices of $G_R(A)$ are named $u_1, \ldots, u_r$. The edge (u_i, u_j) is present precisely when $\{0,1\}^{E_{\rho_i,\rho_j}}$ is in A. The vertex u_i is labelled by the subset of the proper divisors of $|\rho_i|$ equal to $\{k : \{0,1\}^{E_{\rho_i,k}} \in A\}$.

(ii) The graph $G_C(A)$ has s vertices where s is the number of cycles of σ. The vertices of $G_C(A)$ are named $v_1, \ldots, v_s$. The edge (v_i, v_j) is present precisely when $\{0,1\}^{F_{\sigma_i,\sigma_j}}$ is in A. The vertex v_i is labelled by the subset of the proper divisors of $|\sigma_i|$ equal to $\{k : \{0,1\}^{F_{\sigma_i,k}} \in A\}$.

Example 4.4.2. Let $\rho = \sigma = (12)(34)(56)(78)(9) \in S_9$. Denote the cycles of ρ and σ by $\rho_1 = \sigma_1 = (12)$, $\rho_2 = \sigma_2 = (34)$, $\rho_3 = \sigma_3 = (56)$, $\rho_4 = \sigma_4 = (78)$, and $\rho_5 = \sigma_5 = (9)$. Let $A = \{\{0,1\}^{E_{\rho_1,\rho_2}}, \{0,1\}^{E_{\rho_2,\rho_3}}, \{0,1\}^{E_{\rho_3,\rho_4}}, \{0,1\}^{F_{\sigma_1,\sigma_2}}, \{0,1\}^{F_{\sigma_1,\sigma_2}}\}$. Then Figure 4.1 illustrates the corresponding graphs alongside a matrix with entries partitioned as they might be for an element of $\cap A$. Note that this illustration portrays a matrix in $\{0,1\}^{E_{1,3}}$ rather than $\{0,1\}^{E_{1,4}}$. In fact, a matrix in $\{0,1\}^{E_{\rho_1,\rho_2}}$ is in precisely one of those two sets but we can only illustrate one of the possibilities at a time. The alternative would have the blue and red colours swapped in rows 3 and 4, and the same for the green and yellow colours in rows 3 and 4. There were 5 such decisions to make when creating this illustration and to draw all the possibilities would require 2^5 pictures.

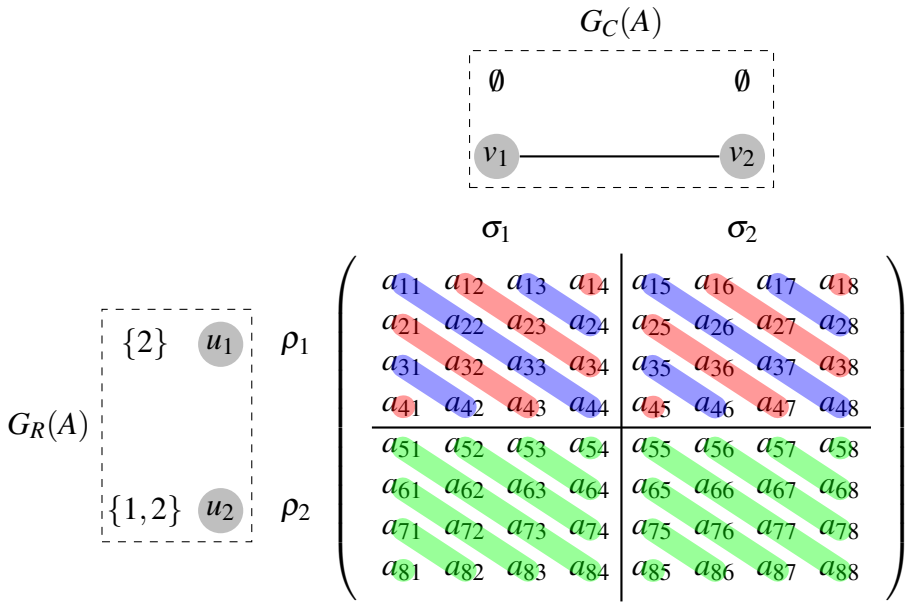


Fig. 4.2 The graph pair corresponding to $\{\{0,1\}^{E_{\rho_1,2}},\{0,1\}^{E_{\rho_2,1}},\{0,1\}^{E_{\rho_2,2}},\{0,1\}^{F_{\sigma_1,\sigma_2}}\}$.

Next we consider an example including vertex labels. Let $\rho=\sigma=(1234)(5678)\in S_8$. Denote the cycles of ρ and σ by $\rho_1=\sigma_1=(1234)$, and $\rho_2=\sigma_2=(5678)$. Let $A=\{\{0,1\}^{E_{\rho_1,2}},\{0,1\}^{E_{\rho_2,1}},\{0,1\}^{E_{\rho_2,2}},\{0,1\}^{F_{\sigma_1,\sigma_2}}\}$. Then Figure 4.2 illustrates the corresponding graphs alongside a matrix with entries partitioned as they might be for an element of $\cap A$.

Next we will define the collection of pairs of graphs $\{(G_R(A),G_C(A)):A\in\mathscr{A}(\rho,\sigma)\}$ without referencing $\mathscr{A}(\rho,\sigma)$. Furthermore we show that the pairs of graphs in this collection is in bijective correspondence with the elements of $\mathscr{A}(\rho,\sigma)$.

Lemma 4.4.3. *Let $(\rho,\sigma)\in S_m\times S_n$. Let $\mathscr{G}(\rho,\sigma)$ denote the collection of all pairs of graphs (G_R,G_C) satisfying the following properties.*

(i) *The number of vertices of G_R is r, the number of distinct cycles of ρ.*

(ii) *The number of vertices of G_C is s, the number of distinct cycles of σ.*

(iii) *There exists a partition $\{p_i\subseteq V(G_R):\exists j\in\mathbf{r}$ such that $i=|\rho_j|\}$ of the vertices of G_R. Those p_i which exist are defined by $u_j\in p_i$ if and only if $|\rho_j|=i$. The labels of the vertices in p_i must be subsets of the divisors of i. Furthermore there are no edges (x,y) in G_R where $x\in p_i$, $y\in p_j$ and $i\neq j$.*

(iv) *There exists a partition $\{q_i\subseteq V(G_C):\exists j\in\mathbf{s}$ such that $i=|\sigma_j|\}$ of the vertices of G_C. Those q_i which exist are defined by $v_j\in q_i$ if and only if $|\sigma_j|=i$. The labels of the vertices*

in q_i must be subsets of the divisors of i. Furthermore there are no edges (x,y) in G_C where $x \in q_i$, $y \in q_j$ and $i \neq j$.

Then the map $\mathscr{A}(\rho,\sigma) \to \mathscr{G}(\rho,\sigma)$ defined by $A \mapsto (G_R(A), G_C(A))$ is a bijection.

Proof. We begin by showing that $A \mapsto (G_R(A), G_C(A))$ is injective. Assume that $A, B \subseteq \mathscr{A}(\rho,\sigma)$ such that $A \neq B$. Then, without loss of generality, assume that there is an element a of A such that a is not an element of B. Either a is of type (i), (ii), (iii), or (iv) of the types of elements in $\mathscr{A}(\rho,\sigma)$ defined at the start of this section. If $a = \{0,1\}^{E_{\rho_i,\rho_j}}$ then (u_i,u_j) is an edge of $G_R(A)$ but not of $G_R(B)$. If $a = \{0,1\}^{F_{\sigma_i,\sigma_j}}$ then (v_i,v_j) is an edge of $G_C(A)$ but not of $G_C(B)$. If $a = \{0,1\}^{E_{\rho_i,k}}$ then k is in the label of the vertex u_i in $G_R(A)$ but not so in $G_R(B)$. If $a = \{0,1\}^{F_{\sigma_i,k}}$ then k is in the label of the vertex v_i in $G_C(A)$ but not so in $G_C(B)$. Therefore in all cases the image of A and B are different therefore the proposed map is injective.

Next we show that $A \mapsto (G_R(A), G_C(A))$ is surjective. Let $(G_R, G_C) \in \mathscr{G}(\rho,\sigma)$. Then define $A \subseteq \mathscr{A}(\rho,\sigma)$ as follows. Let $\{0,1\}^{E_{\rho_i,\rho_j}}$ be in A if (u_i,u_j) is an edge of G_R. Let $\{0,1\}^{F_{\sigma_i,\sigma_j}}$ be in A if (v_i,v_j) is an edge of G_C. Let $\{0,1\}^{E_{\rho_i,k}}$ be in A if k is in the label of $u_i \in G_R$. Let $\{0,1\}^{F_{\sigma_i,k}}$ be in A if k is in the label of $v_i \in G_R$. Then $(G_R(A), G_C(A))$ is equal to (G_R, G_C). Therefore the proposed map is surjective. □

This representation is useful to us because the properties of subsets of $\mathscr{A}(\rho,\sigma)$ which are important will be easier to describe with terminology from graph theory. Some such properties will be part of a (sufficient but not necessary) condition for $A, B \subseteq \mathscr{A}(\rho,\sigma)$ to satisfy $|\cap A| = |\cap B|$. Determining this condition is the focus of Lemma 4.4.5 and Lemma 4.4.6. Recall our convention that ρ has r cycles denoted $\rho_1,\ldots,\rho_r$ and that $G_R(A)$ has r vertices denoted $u_1,\ldots,u_r$. Moreover, u_i corresponds to ρ_i in the sense that $(u_i,u_j) \in G_R(A)$ if and only if $\{0,1\}^{E_{\rho_i,\rho_j}} \in A$. Similarly, our convention is that σ has s cycles denoted $\sigma_1,\ldots,\sigma_s$ and that $G_C(A)$ has s vertices denoted $v_1,\ldots,v_r$. Moreover v_i corresponds to σ_i in the sense that $(v_i,v_j) \in G_C(A)$ if and only if $\{0,1\}^{F_{\sigma_i,\sigma_j}} \in A$. We will be analysing each subset of $\mathscr{A}(\rho,\sigma)$ based on the connected components of the associated pair of graphs. We now create notation to refer to these components, their vertices, and the corresponding cycles of ρ and σ.

We define K_R to be the map which takes a subset A of $\mathscr{A}(\rho,\sigma)$ and returns the connected components of $G_R(A)$. We will let $\mu(A)$ be the number of connected components of $G_R(A)$. We will denote this partition by $K_R(A) = \{K_{R,1}(A),\ldots,K_{R,\mu(A)}(A)\}$. The size of the connected component $K_{R,i}(A)$ will be denoted by $r_i(A)$ and the vertices will be denoted by $\{u_{i,1},\ldots,u_{i,r_i(A)}\}$. For $i \in \mu(\mathbf{A})$ and $j \in \mathbf{r_i}(\mathbf{A})$ we define $\rho_{i,j}$ to be the cycle ρ_k of ρ such that $u_{i,j} = u_k$.

Similarly, we define K_C to be the map which sends a subset A of $\mathscr{A}(\rho,\sigma)$ to the connected components of $G_C(A)$. We will let $\nu(A)$ be the number of connected components of $G_C(A)$. We will denote this partition by $K_C(A) = \{K_{C,1}(A),\ldots,K_{C,\nu(A)}(A)\}$. The size of the

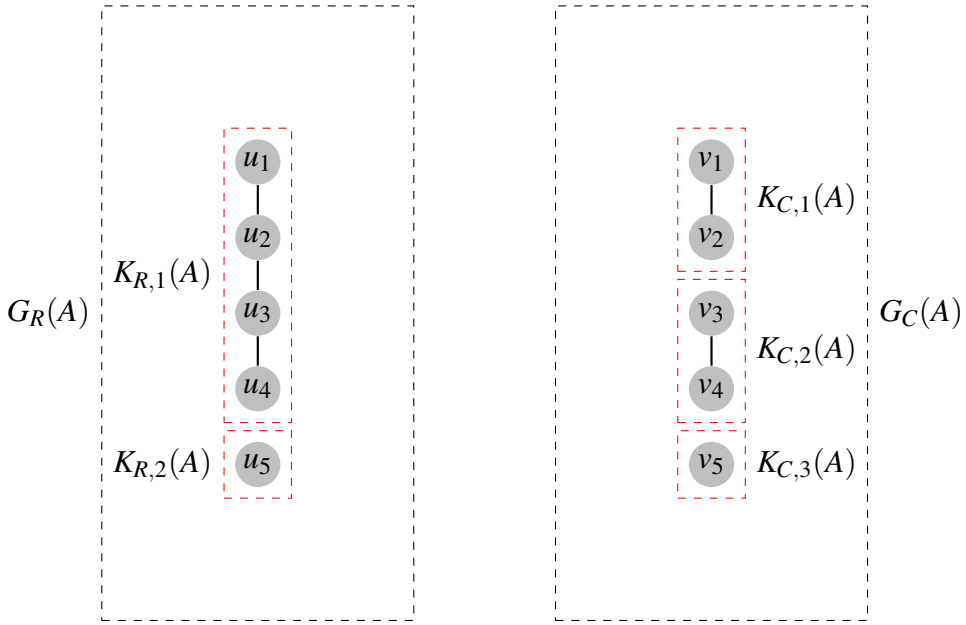


Fig. 4.3 The properties of the graph pair corresponding to $\{\{0,1\}^{E_{\rho_1,\rho_2}},\{0,1\}^{E_{\rho_2,\rho_3}},\{0,1\}^{E_{\rho_3,\rho_4}},\{0,1\}^{F_{\sigma_1,\sigma_2}},\{0,1\}^{F_{\sigma_3,\sigma_4}}\}$.

connected component $K_{C,i}(A)$ will be denoted by $s_i(A)$ and the vertices will be denoted by $\{v_{i,1},\ldots,v_{i,s_i(A)}\}$. For $i \in \nu(\mathbf{A})$ and $j \in \mathbf{s_i}(\mathbf{A})$ we define $\sigma_{i,j}$ to be the cycle σ_k of σ such that $v_{i,j} = v_k$.

Next we create notation to refer to the labels of vertices and related properties for each component of the graphs $G_R(A)$ and $G_C(A)$. We define $L_A : V(G_R(A)) \cup V(G_C(A)) \to \mathscr{P}(\mathbb{N})$ such that the image of a vertex of $G_R(A)$ or $G_C(A)$ is its label. We let $\delta_R(A)$ denote the tuple of natural numbers $(\delta_{R,1}(A),\ldots,\delta_{R,\mu(A)}(A))$ such that $\delta_{R,i}(A)$ is the greatest common divisor of the labels of vertices in the connected component $K_{R,i}(A)$. If all vertex labels in the connected component $K_{R,i}(A)$ are empty then $\delta_{R,i}(A)$ is set to equal $|\rho_{i,1}|$. Similarly, we let $\delta_C(A)$ denote the tuple of natural numbers $(\delta_{C,1}(A),\ldots,\delta_{C,\nu(A)}(A))$ such that $\delta_{C,i}(A)$ is the greatest common divisor of the labels of vertices in the connected component $K_{C,i}(A)$. If all vertex labels in the connected component $K_{C,i}(A)$ are empty then $\delta_{C,i}(A)$ is set to equal $|\sigma_{i,1}|$.

Example 4.4.4. Let us recall the graphs in Example 4.4.2. Then Figure 4.3 and Figure 4.4 show the components and Figure 4.2 also shows the greatest common divisors of the labels for each component.

For Lemma 4.4.5 we must recall the definition of the *transitive closure* of a relation. First we define the *transitive extension* of a relation $R \subseteq X \times X$ to be the relation which contains both R and the set

$$\{(x,y) : \exists z \in X \text{ such that } (x,z),(z,y) \in R\}.$$

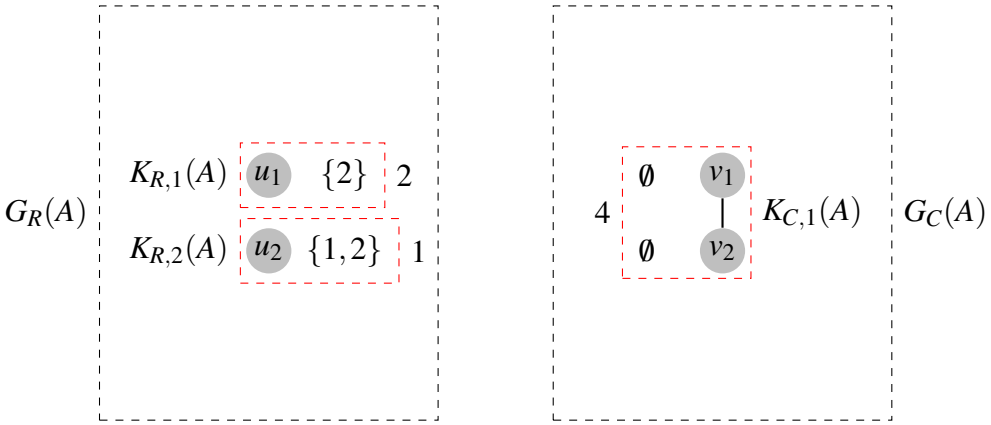


Fig. 4.4 The properties of the graph pair corresponding to $\{\{0,1\}^{E_{\rho_1,2}},\{0,1\}^{E_{\rho_2,1}},\{0,1\}^{E_{\rho_2,2}},\{0,1\}^{F_{\sigma_1,\sigma_2}}\}$. The number labelling the component $K_{X,i}(A)$ is the value of $\delta_{X,i}(A)$.

If R_1 denotes the transitive extension of $R \subseteq X \times X$, and R_{i+1} denotes the transitive extension of R_i then the transitive closure of R, sometimes denoted R^∞, is the set union of $R_1, R_2, \ldots$. For our purposes, the *transitive closure* of the simple graph G is a graph with the same vertex set $V(G)$ as G and edge set equal to the transitive closure of $E(G)$ when it is viewed as a relation on $V(G)$, and we do not include any loops. That is to say, the transitive closure of G has vertex set $V(G)$ and edge set equal to

$$(E(G))^\infty \backslash \{(x,x) : x \in V(G)\}$$

i.e. the transitive closure of $E(G) \subseteq V(G) \times V(G)$ minus the reflexive pairs in $V(G) \times V(G)$. Lemma 4.4.5 provides a sufficient (but not necessary) condition for $A, B \subseteq \mathscr{A}(\rho,\sigma)$ to satisfy $\cap A = \cap B$ and will be important to defining a sufficient condition for the more general statement $|\cap A| = |\cap B|$ to hold.

Lemma 4.4.5. *Let $m,n \in \mathbb{N}$ and let $(\rho,\sigma) \in S_m \times S_n$. For all $A \subseteq \mathscr{A}(\rho,\sigma)$ we will define $\bar{A}$ to be the subset of $\mathscr{A}(\rho,\sigma)$ such that:*

(i) *$A \subseteq \bar{A}$;*

(ii) *The graph $G_R(\bar{A})$ has edge set equal to that of the transitive closure of $G_R(A)$ and the graph $G_C(\bar{A})$ has edge set equal to that of the transitive closure of $G_C(A)$;*

(iii) *For all pairs (i,j) such that $1 \leq i \leq r$ and $1 \leq j \leq r_i(A)$ we have that*

$$L_{\bar{A}}(u_{i,j}) = \{k : \delta_{R,i} \text{ divides } k,\ k \text{ divides } |\rho_{i,j}|,\ \text{and } k \neq |\rho_{i,j}|\}.$$

Similarly, for all pairs (i, j) *such that* $1 \leq i \leq s$ *and* $1 \leq j \leq s_i(A)$ *we have that*

$$L_{\bar{A}}(v_{i,j}) = \{k : \delta_{C,i} \text{ divides } k,\ k \text{ divides } \left|\sigma_{i,j}\right|,\ \text{and } k \neq \left|\sigma_{i,j}\right|\}.$$

Essentially, to construct $\bar{A}$ *from* A *we have added as many edges as possible to* $G_R(A)$ *and* $G_C(A)$ *without changing the sizes* $r_i(A)$ *and* $s_j(A)$ *(respectively) of the connected components of either graph. We have also added as many elements to the labels of vertices of* $G_R(A)$ *and* $G_C(A)$ *as possible without changing the values of* δ_R *and* δ_C*. We then have that*

$$\cap A = \cap \bar{A}$$

Proof. Since $A \subseteq \bar{A}$ it is clear that $\cap \bar{A} \subseteq \cap A$ so we only need to prove the reverse containment. To do this, we will show that $a \in \bar{A}$ implies $\cap A \subseteq a$. If this is true for all $a \in \bar{A}$ then we will have that $\cap A \subseteq \cap \bar{A}$. Let $a \in \bar{A} \backslash A$. Then one of the following holds

(i) $a \notin A$ corresponds to an edge in $G_R(\bar{A})$ or $G_C(\bar{A})$ which was not an edge of $G_R(A)$ or $G_C(\bar{A})$, respectively; or

(ii) for some $u \in V(G_C(A)) \cup V(G_R(A))$ we have that $a \notin A$ corresponds to an element of the vertex label $L_{\bar{A}}(u)$ which is not an element of the vertex label $L_A(u)$.

In case (i) a corresponds to an edge in the transitive closure of $G_R(A)$ or $G_C(A)$. Assume that a corresponds to an edge in the transitive closure of $G_R(A)$. Recall that

$$\{0,1\}^{E_{\rho_x,\rho_y}} = \bigcup_{i \in \mathrm{dom}(\rho_x)} \bigcup_{j \in \mathrm{dom}(\rho_y)} \{0,1\}^{E_{i,j}}.$$

Assume that a corresponds to $\{0,1\}^{E_{\rho_x,\rho_y}}$. Then there must some ρ_z such that $\{0,1\}^{E_{\rho_x,\rho_z}}$ and $\{0,1\}^{E_{\rho_z,\rho_y}}$ are both in A. Choose $k \in \mathrm{dom}(\rho_z)$. For any i in $\mathrm{dom}(\rho_x)$ and j in $\mathrm{dom}(\rho_y)$ we have

$$\{0,1\}^{E_{i,k}} \subseteq \{0,1\}^{E_{\rho_x,\rho_z}} \in A \text{ and } \{0,1\}^{E_{k,j}} \subseteq \{0,1\}^{E_{\rho_z,\rho_y}} \in A$$

which implies

$$\cap A \subseteq \{0,1\}^{E_{i,k}} \cap \{0,1\}^{E_{k,j}}.$$

Furthermore

$$\{0,1\}^{E_{i,k}} \cap \{0,1\}^{E_{k,j}} \subseteq \{0,1\}^{E_{i,j}}$$

because if row i equals row k and row k equals row j then row i equals row j. Therefore $\cap A \subseteq \{0,1\}^{E_{\rho_x,\rho_y}}$ as required. The case where a corresponds to an edge of the transitive closure of $G_C(A)$ is similar.

In case (ii) assume a corresponds to $\{0,1\}^{E_{\rho_{i,j},k}}$ for some $i \in \mu(\mathbf{A})$ and $j \in \mathbf{r_i}(\mathbf{A})$. By definition we have that $\{0,1\}^{E_{\rho_{i,j},\delta_{R,i}}} \subseteq \{0,1\}^{E_{\rho_{i,j},k}}$ and we will show that $\cap A \subseteq \{0,1\}^{E_{\rho_{i,j},\delta_{R,i}}}$. The greatest common divisor of the labels of vertices in the component $K_{R,i}$ would not be equal to $\delta_{R,i}$. Therefore there must be

$$\{0,1\}^{E_{\rho_{i,j_1},k_1}}, \ldots, \{0,1\}^{E_{\rho_{i,j_p},k_p}} \in A$$

such that the greatest common divisor of $k_1, \ldots, k_p$ is $\delta_{R,i}$. Let $z \in \mathrm{dom}(\rho_{i,j})$ and let $z_1, \ldots z_p$ be in $\mathrm{dom}(\rho_{i,j_1}), \ldots, \mathrm{dom}(\rho_{i,j_p})$, respectively, such that $\{0,1\}^{E_{z,z_1}}, \ldots, \{0,1\}^{E_{z,z_p}}$ are each contained in elements of A. Then we have

$$\{0,1\}^{E_{z,z_l}} \cap \{0,1\}^{E_{\rho_{i,l},k_l}} \subseteq \{0,1\}^{E_{\rho i,j,k_l}}$$

for all $1 \leq l \leq p$ and it follows that

$$\bigcap_{l=1}^{p} \left(\{0,1\}^{E_{z,z_l}} \cap \{0,1\}^{E_{\rho_{i,l},k_l}} \right) \subseteq \bigcap_{l=1}^{p} \{0,1\}^{E_{\rho_{i,j},k_l}}.$$

Finally we note that $\{0,1\}^{E_{\rho_{i,j},\delta_{R,i}}} = \cap_{l=1}^{p} \{0,1\}^{E_{\rho_{i,j},k_l}}$ and therefore

$$\cap A \subseteq \bigcap_{l=1}^{p} \{0,1\}^{E_{\rho_{i,j},k_l}} \subseteq \{0,1\}^{E_{\rho_{i,j},\delta_{R.i}}} \subseteq \{0,1\}^{E_{\rho_{i,j},k}}$$

as required. □

By Lemma 4.4.5 we may partition $\mathscr{P}(\mathscr{A}(\rho,\sigma))$ into classes of the form

$$[A] = \{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}.$$

However we can do even better and partition $\mathscr{A}(\rho,\sigma)$ into larger classes where all elements have the same size of intersection, even if their intersections are not equal. To do this we define an action of $S_m \times S_n$ on $\mathscr{P}(\mathscr{A}(\rho,\sigma))$ such that elements of an orbit have the same cardinality of intersection. In order to do this, we first define an action of $S_m \times S_n$ on $\mathscr{A}(\rho,\sigma)$ which will induce the aforementioned action on the power set. For each $a \in \mathscr{A}(\rho,\sigma)$ and each $(\mu,\nu) \in S_m \times S_n$ we let

$$a \cdot (\mu,\nu) = \{f \cdot (\mu,\nu) : f \in a\}$$

where $f \cdot (\mu, \nu)$ is the usual action defined by $(i,j)f \cdot (\mu,\nu) = (i\mu^{-1}, j\nu^{-1})f$ for all (i,j) in $\mathbf{m} \times \mathbf{n}$. We extend this map to a map on $\mathscr{P}(\mathscr{A}(\rho,\sigma))$ by

$$A \cdot (\mu,\nu) = \{a \cdot (\mu,\nu) : a \in A\}.$$

We note that $a \mapsto \{f \cdot (\mu,\nu) : f \in a\}$ is a bijection from $\mathscr{A}(\rho,\sigma)$ to $\mathscr{A}(\mu^{-1}\rho\mu, \nu^{-1}\sigma\nu)$ and that $A \mapsto \{a(\mu,\nu) : a \in A\}$ is a bijection from $\mathscr{P}(\mathscr{A}(\rho,\sigma))$ to $\mathscr{P}(\mathscr{A}(\mu^{-1}\rho\mu, \nu^{-1}\sigma\nu))$. If we restrict our attention to the normalizer $N_{S_m \times S_n}((\rho,\sigma))$ of (ρ,σ) in $S_m \times S_n$ then we obtain an action on $\mathscr{P}(\mathscr{A}(\rho,\sigma))$. The normalizer of a single element of a group is just the centralizer of the element so we will in fact consider the action of the centralizer $C_{S_m \times S_n}((\rho,\sigma))$. The claim we will now prove is that two elements A, B in the same orbit of this action satisfy $|\cap A| = |\cap B|$.

Lemma 4.4.6. *Let $m,n \in \mathbb{N}$ and let $(\rho,\sigma) \in S_m \times S_n$. Let $(\mu,\nu) \in C_{S_m \times S_n}((\rho,\sigma))$ then*

$$a \cdot (\mu,\nu) = \{f \cdot (\mu,\nu) : f \in a\}$$

defines an action of $C_{S_m \times S_n}((\rho,\sigma))$ on $\mathscr{A}(\rho,\sigma)$. The corresponding action on $\mathscr{P}(\mathscr{A}(\rho,\sigma))$ satisfies

$$|\cap A| = |\cap (A \cdot (\mu,\nu))|.$$

Proof. It is clear that $a \mapsto \{f \cdot (\mu,\nu) : f \in a\}$ is a bijection however we must show that the image is an element of $\mathscr{A}(\rho,\sigma)$. This follows immediately from the fact that $\{f \cdot (\mu,\nu) : f \in \{0,1\}^{E_{x,y}}\} = \{0,1\}^{E_{x\mu,y\nu}}$ since a is either equal to some $\{0,1\}^{E_{x,y}}$ or equal to the union of several. The other requirements for this to be an action follow from the fact that $(x,y)f \mapsto (x\mu^{-1}, y\nu^{-1})f$ defines an action of $S_m \times S_n$ on binary matrices. It also follows immediately that $|a| = |a \cdot (\mu,\nu)|$. Since $\cap A \cdot (\mu,\nu) = \{f \cdot (\mu,\nu) : f \in \cap A\}$ we have that $|\cap A| = |\cap A \cdot (\mu,\nu)|$ as required. □

Lemma 4.4.6 shows that there is a partition of $\mathscr{P}(\mathscr{A}(\rho,\sigma))$, where elements in the same class have the same size of intersection, defined by

$$[A] = \{C \subseteq \mathscr{A}(\rho,\sigma) : \exists B \subseteq \mathscr{A}(\rho,\sigma); \bar{B} = \bar{A}; C \in B^{C_{S_m \times S_n}((\rho,\sigma))}\}.$$

This can equivalently be defined as

$$[A] = \{B \subseteq \mathscr{A}(\rho,\sigma) : \exists (\mu,\nu) \in C_{S_m \times S_n}((\rho,\sigma)) : \bar{A} = \bar{B} \cdot (\mu,\nu)\}.$$

Herein we will denote this partition by $P(\rho,\sigma)$. We note that the action of an element of $C_{S_m \times S_n}((\rho,\sigma))$ on a subset of $\mathscr{A}(\rho,\sigma)$ corresponds to a permutation of vertices of $G_R(A)$

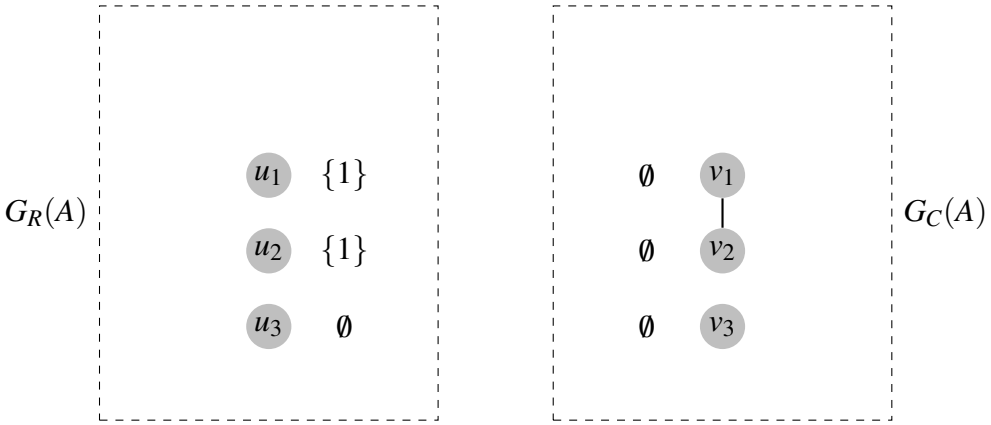


Fig. 4.5 The graph pair $(G_R(A), G_C(A))$ relating to Example 4.4.7

which can send a vertex to any other which corresponds to a cycle of the same length, and the same for $G_C(A)$. This is illustrated in Example 4.4.7.

Example 4.4.7. Consider $\rho = \sigma = (1,2)(3,4)(5,6) \in S_6$ and the subset A of $\mathscr{A}(\rho,\sigma)$ which corresponds to the graph pair showing in Figure 4.5. Note that A is equal to $\bar{A}$. However if $\sigma_1 = (1,2), \sigma_2 = (3,4), \sigma_3 = (5,6)$ then the action of $(1_{S_6}, (1,5)(2,6)) \in C_{S_6 \times S_6}(\rho,\sigma)$ swaps the vertices v_1 and v_3 in $G_C(A)$ so that the connect component of size 2 becomes $\{v_2, v_3\}$. Let $B = A \cdot (1_{S_6}, (1,5)(2,6))$ then $B \neq A$ and $\bar{B} = B$. This shows how $[A]_{P(\rho,\sigma}$ is in general a coarser equivalence than $X \sim Y$ if and only if $\bar{X} = \bar{Y}$. The orbit of A in $C_{S_6 \times S_6}(\rho,\sigma)$ has size 9 since we can send u_3 to u_1, u_2, or u_3, and independently send v_3 to v_1, v_2, or v_3. Note that swapping u_1 and u_2 fixes A and the same is true of swapping v_1 and v_2.

Let $\sim$ denote the equivalence on $\mathscr{A}(\rho,\sigma)$ such that $A \sim B$ if and only if $\bar{A} = \bar{B}$. It will be important later to understand the decomposition of the class $[A]_{P(\rho,\sigma)}$ into a union of classes of $\sim$. The equivalence $\sim$ is clearly coarser than the equivalence corresponding to $P(\rho,\sigma)$ by the definition of the later including the definition of the former. Thus a class $[A]_{P(\rho,\sigma)}$ is a union of classes of $\sim$. These classes are permuted by the action of $C_{S_m \times S_n}((\rho,\sigma))$, whose action is most intuitively seen as permuting the vertices of the corresponding graph pairs (any two vertices of one of the graphs which correspond to cycles of equal length can be swapped). This intuition is illustrated in Example 4.4.7 when we describe the orbit of $A = \bar{A}$.

In order to apply the inclusion-exclusion principle as in Equation 4.4 with terms collected according to the partition $P(\rho,\sigma)$, we must be able to do two things for each $p \in P(\rho,\sigma)$. First, we must be able to take some $A \in p$ and calculate the cardinality of $\cap A$, which is the same value for all elements of p. Second, we must determine the value k_p from Equation (4.4), which is the sum over elements of p where we add 1 for an element of the power set of $\mathscr{A}(\rho,\sigma)$ with

odd cardinality and -1 for an element with even cardinality. We will now tackle the former issue in Section 4.4.2 and the latter in Section 4.4.3.

4.4.2 Determining the cardinality of matrix set intersections

$$
\begin{array}{c|c}
 & \begin{array}{cccc|cccc|c} \sigma_1 & & & & \sigma_2 & & & & \sigma_3 \end{array} \\
\rho_1 & \left(\begin{array}{cccc|cccc|c}
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ \hline
0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ \hline
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1
\end{array}\right) \\
\begin{array}{c} \\ \rho_2 \\ \\ \\ \\ \rho_3 \end{array} &
\end{array}
$$

	$j=1$	$j=2$	$j=3$
$i=1$	4	4	1
$i=2$	2	1	1
$i=3$	1	1	1

Fig. 4.6 An example of a matrix f and a table showing the values $\lambda_f(i,j)$ of the corresponding function λ_f.

In this section we will show how to calculate the cardinality of the intersection of any subset A of $\mathscr{A}(\rho,\sigma)$. In order to do this we will partition the elements of $\cap A$ in a certain way and take the sum of the sizes of all of the classes of that partition. In order to create this partition we define the map

$$\lambda_f : \mu(\mathbf{A}) \times \nu(\mathbf{A}) \to \mathbb{N}$$

by setting $\lambda_f(i,j)$ to equal the row period of the sub-matrix $f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})}$ of f for all $(i,j) \in \mu(\mathbf{A}) \times \nu(\mathbf{A})$. See Figure 4.6 for an example of such a function. We denote the collection of all such functions for matrices in $\cap A$ by

$$\Lambda(A) = \{\lambda_f : f \in \cap A\}.$$

Each of the functions in $\Lambda(A)$ corresponds to a class of the following partition of $\cap A$:

$$Q_A = \{\{f \in \cap A : \lambda_f = \lambda\} : \lambda \in \Lambda(A)\}.$$

In other words, Q_A is the set of inverse images of the map $f \mapsto \lambda_f$. This partition has elements $f,g \in \cap A$ in the same class if and only if $\lambda_f = \lambda_g$. We have $f \in \{0,1\}^{E_{\rho_{i,1},\delta_{R,i}(A)}}$ for all $i \in \mu(\mathbf{A})$ and $f \in \{0,1\}^{F_{\sigma_{j,1},\delta_{C,j}(A)}}$ for all $j \in \nu(\mathbf{A})$. Applying Lemma 4.3.20 then tells us that the row

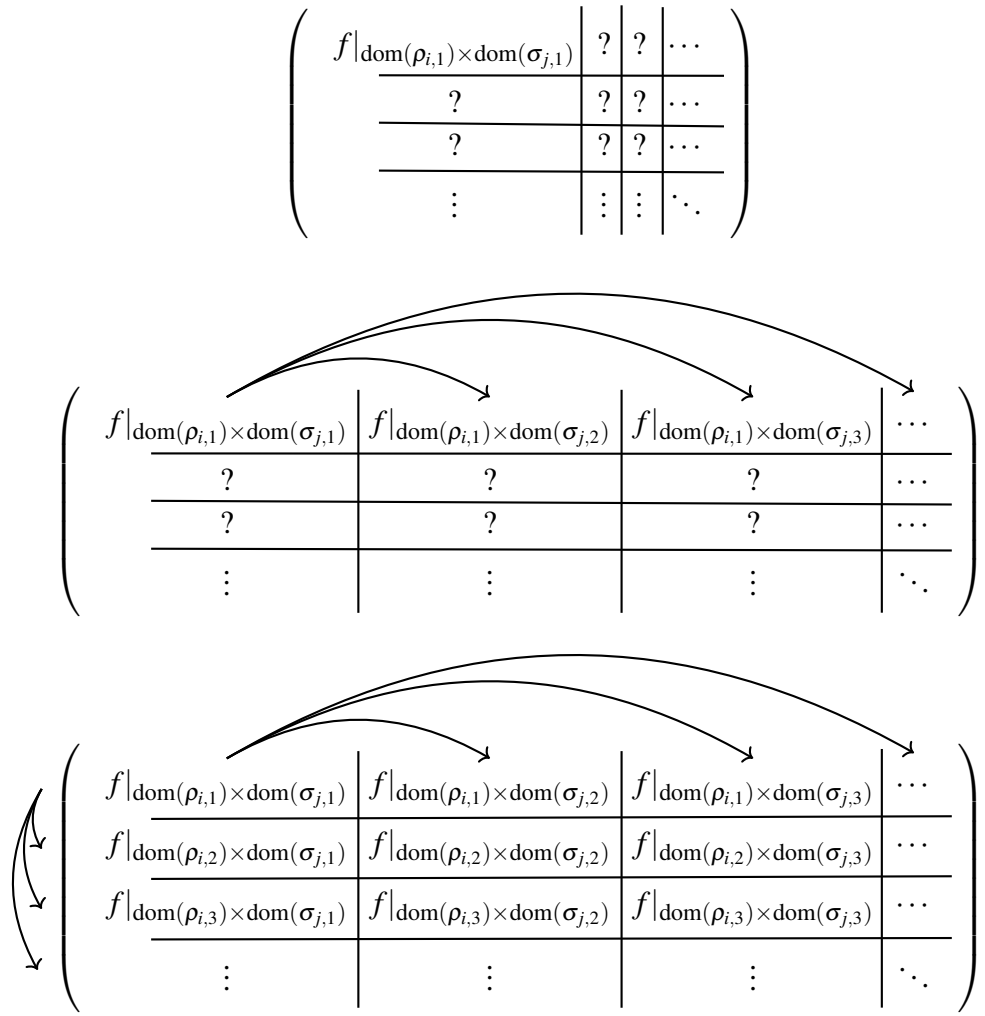


Fig. 4.7 Consider the a matrix f in some $\cap A$ and restrict our attention to the sub-matrix corresponding to some components $K_{R,i}(A)$ and $K_{C,j}(A)$. Then this figure shows how a matrix corresponds to: a choice of $f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})}$; followed by a choice of which rows with indices in $\mathrm{dom}(\rho_{i,2}), \mathrm{dom}(\rho_{i,2}), \ldots$ are equal to some fixed row with index in $\mathrm{dom}(\rho_{i,1})$; and finally a choice of which columns with indices in $\mathrm{dom}(\sigma_{j,2}), \mathrm{dom}(\sigma_{j,2}), \ldots$ are equal to some fixed column with index in $\mathrm{dom}(\sigma_{j,1})$.

period of $f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})}$ divides

$$\gcd(\delta_{R,i}(A), \delta_{C,j}(A))$$

for all $(i,j) \in \mu(\mathbf{A}) \times \nu(\mathbf{A})$. It may be seen that $\Lambda(A)$ can be equivalently defined as the set

$$\{\lambda : \mu(\mathbf{A}) \times \nu(\mathbf{A}) \to \mathbb{N} : \lambda(i,j) | \gcd(\delta_{R,i}(A), \delta_{C,j}(A)(A)) \text{ for all } (i,j) \in \mu(\mathbf{A}) \times \nu(\mathbf{A})\}.$$

This is helpful as it makes the classes of Q_A easy to count or iterate through. The end result of this section, Theorem 4.4.15, is akin to Theorem 4.4.8.

Theorem 4.4.8. *Let $(\rho,\sigma) \in S_m \times S_n$ and let A be a subset of $\mathscr{A}(\rho,\sigma)$. Then*

$$\left|\bigcap_{a\in A} a\right| = \sum_{\lambda\in\Lambda(A)} \Omega(A,\lambda),$$

where Ω maps $A \in \mathscr{A}(\rho,\sigma)$ and $\lambda \in \Lambda(A)$ to the size of $\{f \in \cap A : \lambda_f = \lambda\}$.

Before we can give a proof of Theorem 4.4.8, we describe $\Omega(A,\lambda)$ in a combinatorial manner. There are three pieces to this combinatorial puzzle. First, the matrices in $\{f \in \cap A : \lambda_f = \lambda\}$ can have a specific number of possibilities for each sub-matrices on the domains

$$\{\mathrm{dom}(\rho_{i,1}) \times \mathrm{dom}(\sigma_{j,1}) : (i,j) \in \mu(\mathbf{A}) \times \nu(\mathbf{A})\}.$$

The choice for each sub-matrix is independent of all the others. Second, for each $i \in \mu(\mathbf{A})$ and $1 < i' \leq r_i(A)$ there is $x \in \mathrm{dom}(\rho_{i,1})$ and $y \in \mathrm{dom}(\rho_{i,i'})$ such that row x equals row y. Although x and y are not unique with this property, we can describe which choices of $x \in \mathrm{dom}(\rho_{i,1})$ and $y \in \mathrm{dom}(\rho_{i,i'})$ correspond to distinct outcomes. The final piece is analogous to the second but concerns columns instead of rows. For each $j \in \nu(\mathbf{A})$ and $1 < j' \leq s_j$ there is $x \in \mathrm{dom}(\sigma_{j,1})$ and $y \in \mathrm{dom}(\sigma_{j,j'})$ such that column x equals column y. Again, we can describe which choices of $x \in \mathrm{dom}(\sigma_{j,1})$ and $y \in \mathrm{dom}(\sigma_{j,j'})$ correspond to distinct outcomes. The number $\Omega(A,\lambda)$ is just the product of the number of possibilities for the aforementioned sub-matrices, times the number of choices of $x \in \mathrm{dom}(\rho_{i,1})$ and $y \in \mathrm{dom}(\rho_{i,i'})$ for each $i \in \mu(\mathbf{A})$ and $1 < i' \leq r_i(A)$ which lead to distinct outcomes, times the number of choices of $x \in \mathrm{dom}(\sigma_{j,1})$ and $y \in \mathrm{dom}(\sigma_{j,j'})$ for each $j \in \nu(\mathbf{A})$ and $1 < j' \leq s_j$ which lead to distinct outcomes. Figure 4.7 is an illustration of how to imagine a matrix being formed through these three choices.

We start by working on the aforementioned first piece of the puzzle. That is to say, knowing how many possibilities there are for the sub-matrix $f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})}$ given the value $\lambda_f(i,j)$ of the row period of this sub-matrix.

Lemma 4.4.9. *Let $(\rho,\sigma) \in S_m \times S_n$. Let ρ_i be a cycle of ρ and let σ_j be a cycle of σ. Let q be a divisor of* $\gcd(|\rho_i|,|\sigma_j|)$. *Then define $\zeta(\rho,\sigma,\rho_i,\sigma_j,q)$ to be the collection*

$$\{f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)} : f \in \{0,1\}^{R_{\rho,\sigma}} \text{ and the row period of } f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)} \text{ is } q\}.$$

of all sub-matrices on the domain $\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)$ *of matrices in $\{0,1\}^{R_{\rho,\sigma}}$, where the sub-matrix has row period equal to q. Then:*

(i) *Equivalently $\zeta(\rho,\sigma,\rho_i,\sigma_j,q)$ can be defined as the collection*

$$\{f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)} : f \in \{0,1\}^{R_{\rho,\sigma}} \text{ and the column period of } f|_{\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)} \text{ is } q\}.$$

(ii) *We can calculate the size of $\zeta(\rho,\sigma,\rho_i,\sigma_j,q)$ by:*

$$\left|\zeta(\rho,\sigma,\rho_i,\sigma_j,q)\right| = 2^q - \sum_{q'\in\mathrm{div}(q)} \left|\zeta(\rho,\sigma,\rho_i,\sigma_j,q')\right|.$$

where $\mathrm{div}(q)$ *is the set of proper divisors of q.*

(iii) *We can determine all the values $\{\left|\zeta(\rho,\sigma,\rho_i,\sigma_j,q)\right| : q$ divides $\gcd(|\rho_i|,|\sigma_j|)\}$ recursively using $\left|\zeta(\rho,\sigma,\rho_i,\sigma_j,1)\right| = 2$ as a start point.*

(iv) *The size $\left|\zeta(\rho,\sigma,\rho_i,\sigma_j,q)\right|$ of this collection of matrices depends only on q and not on the specific ρ,σ,ρ_i and σ_j other than the restriction that q must be a divisor of both $|\rho_i|$ and $|\sigma_j|$.*

Proof. (i) This follows from Lemma 4.3.16.

(ii) We first show that $\left|\cup_{q'|q}\zeta(\rho,\sigma,\rho_i,\sigma_j,q')\right| = 2^q$. To do this we partition the elements of $\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)$ so that $(x_1,y_1),(x_2,y_2)$ are in the same part if $(x_1,y_1)f$ is equal to $(x_2,y_2)f$. Let $(x,y) \in \mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)$. Then we have that $(x,y)f = (x\rho^z, y\sigma^z)f$ for all $z \in \mathbb{Z}$. We also have that $(x,y)f = (x\rho^{kq},y)f$ for all $q \in \mathbb{Z}$. It follows that $(x,y)f = (x\rho^{z+kq}, y\sigma^z)f$ for all $k,z \in \mathbb{Z}$. The partition

$$\{\{(x\rho^{z+kq+a}, y\sigma^z) : k,z \in \mathbb{Z}\} : 0 \leq a \leq q-1\}$$

is the partition we seek. There are q many classes and either all entries of f with indices in a certain class are 0 or they are 1. Therefore there are 2^q possibilities.

Now since a matrix cannot have two different row periods, the union

$$\bigcup_{q'|q} \zeta(\rho,\sigma,\rho_i,\sigma_j,q')$$

is a disjoint union. Therefore

$$|\bigcup_{q'|q} \zeta(\rho,\sigma,\rho_i,\sigma_j,q')| = \sum_{q'|q} |\zeta(\rho,\sigma,\rho_i,\sigma_j,q')|$$

and since this expression is equal to 2^q the result follows immediately.

(iii) We notice that $\left|\zeta(\rho,\sigma,\rho_i,\sigma_j,1)\right| = 2$ as the only possibilities in that case are the all zero or the all one matrices. Then we can use this value to determine $\left|\zeta(\rho,\sigma,\rho_i,\sigma_j,q)\right|$ for those q which are prime. Then we can additionally use those values to determine the q which are the product of two primes, and so on.

(iv) It is clear from (iii) that $\left|\zeta(\rho,\sigma,\rho_i,\sigma_j,q)\right|$ does not depend on ρ,σ,ρ_i, or σ_j other than the fact that q must be a divisor of both $|\rho_i|$ and $|\sigma_j|$.

□

We will write $\omega(q)$ to denote the value of $\left|\zeta(\rho,\sigma,\rho_i,\sigma_j,q)\right|$ when q is a divisor of both $|\rho_i|$ and $|\sigma_j|$. We define ω because Lemma 4.4.9 part (iv) demonstrated that q is the cardinality of the set returned by ζ is entirely dependent on q in this situation. We note that if $f \in \cap A$ then the row period of $f|_{\mathrm{dom}(\rho_{i,j})\times\mathbf{n}}$ is equal to $\mathrm{lcm}\{\lambda(i,k) : k \in \nu(\mathbf{A})\}$ and the column period of $f|_{\mathbf{m}\times\mathrm{dom}(\sigma_{i,j})}$ is equal to $\mathrm{lcm}\{\lambda(k,i) : k \in \mu(\mathbf{A})\}$. To help us conveniently refer to these values we will write

$$\lambda(i,*) = \mathrm{lcm}\{\lambda(i,k) : k \in \nu(\mathbf{A})\}$$

and

$$\lambda(*,i) = \mathrm{lcm}\{\lambda(k,i) : k \in \mu(\mathbf{A})\}.$$

We also note that if $f \in \cap A$ then $\lambda_f(i,*)$ divides $\delta_{R,i}(A)$ and $\lambda_f(*,i)$ divides $\delta_{C,i}(A)$.

The next lemma will also be important to the first piece of the combinatorial problem. It involves the correspondence between the sub-matrices of a matrix f of the form $f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})}$, the function λ_f, and membership of certain matrix sets of the form $\{0,1\}^{E_{\rho_{i,i'},k}}$ or $\{0,1\}^{F_{\sigma_{j,j'},k}}$.

Lemma 4.4.10. *Let $(\rho,\sigma) \in S_m \times S_n$. Let $A \subseteq \mathscr{A}(\rho,\sigma)$. Let $\lambda \in \Lambda(A)$. Let $f \in \cap A$. For all $(i,j) \in \mu(\mathbf{A}) \times \nu(\mathbf{A})$ assume that $f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})} \in \zeta(\rho,\sigma,\rho_{i,1},\sigma_{j,1},\lambda(i,j))$. Then the following are true*

(i) $\lambda_f = \lambda$;

(ii) $f \in \{0,1\}^{E_{\rho_{i,j},\lambda(i,*)}}$ *for all* $i \in \mu(\mathbf{A})$ *and* $j \in \mathbf{r_i}(\mathbf{A})$;

(iii) $f \in \{0,1\}^{F_{\sigma_{i,j},\lambda(*,i)}}$ *for all* $i \in \nu(\mathbf{A})$ *and* $j \in \mathbf{s_i}(\mathbf{A})$.

Proof. First, $f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})} \in \zeta(\rho,\sigma,\rho_{i,1},\sigma_{j,1},\lambda(i,j))$ means that $f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})}$ has row period $\lambda(i,j)$. Therefore (i) follows from the definition of λ_f. Next, note that $f \in \{0,1\}^{E_{\rho_{i,j},\lambda(i,*)}}$ if and only if the row period of $f|_{\mathrm{dom}(\rho_{i,j})\times\mathbf{n}}$ divides $\lambda(i,*)$ by Lemma 4.3.20. The row period of $f|_{\mathrm{dom}(\rho_{i,j})\times\mathbf{n}}$ is $\lambda_f(i,*)$ by definition so, since (i) is also true, (ii) must be true. Part (iii) is true by an analogous argument to that used for part (ii). □

The second piece in our combinatorial problem of determining $\Omega(A,\lambda)$ starts with us investigating the implications of a matrix being in $\{0,1\}^{E_{\rho_i,\rho_j}}$, i.e. having some row with index in $\mathrm{dom}(\rho_i)$ equal to some row with index in $\mathrm{dom}(\rho_j)$. The corresponding situation with columns relates to the third piece of our combinatorial problem. It is equally important but entirely analogous.

Lemma 4.4.11. *Let* $(\rho,\sigma) \in S_m \times S_n$. *Let* $A \subseteq \mathscr{A}(\rho,\sigma)$. *Let* $i \in \mu(\mathbf{A})$ *such that the number of vertices of the connected component* $K_{R,i}(A)$ *of* $G_R(A)$ *is greater than one. Let* $1 \leq j_1 < j_2 \leq r_i(A)$, *so that* u_{i,j_1}, u_{i,j_2} *are vertices of* $K_{R,i}(A)$. *Let* $x_1 \in \mathrm{dom}(\rho_{i,j_1})$ *and* $x_2 \in \mathrm{dom}(\rho_{i,j_2})$, *so that* x_1, x_2 *are in the domains of the cycles corresponding to the vertices* u_{i,j_1}, u_{i,j_2}. *Let* $f \in \{0,1\}^{R_{\rho,\sigma}}$ *be a matrix fixed by* (ρ,σ). *Then* $f \in \{0,1\}^{E_{\rho_{i,j_1},\rho_{i,j_2}}}$ *if and only if there exists* $k \in \lambda_{\mathbf{f}}(\mathbf{i},*)$ *such that* $f \in \{0,1\}^{E_{x_1,x_2'}}$ *where* $x_2' = x_2\rho^k$. *Furthermore this* k *is unique in* $\lambda_{\mathbf{f}}(\mathbf{i},*)$. *An analogous result holds for columns.*

Proof. We begin with the forward implication, by assuming $f \in \{0,1\}^{E_{\rho_{i,j_1},\rho_{i,j_2}}}$. By definition, $f \in \{0,1\}^{E_{\rho_{i,j_1},\rho_{i,j_2}}}$ implies that there is an $a \in \mathrm{dom}(\rho_{i,j_1})$ and $b \in \mathrm{dom}(\rho_{i,j_2})$ such that $f \in \{0,1\}^{E_{a,b}\vee R_{\rho,\sigma}}$. Since $a, x_1 \in \mathrm{dom}(\rho_{i,j_2})$ there is $\alpha \in \mathbb{N}$ such that $x_1\rho^\alpha = a$. Similarly, there is $\beta \in \mathbb{N}$ such that $x_2\rho^{\alpha+\beta} = b$. We note that $E_{a,b} \vee R_{\rho,\sigma}$ is equal to $E_{a\rho^k,b\rho^k} \vee R_{\rho,\sigma}$ for all $k \in \mathbb{Z}$. Therefore we have that:

$$\begin{aligned} E_{a,b} \vee R_{\rho,\sigma} &= E_{x_1\rho^\alpha, x_2\rho^{\alpha+\beta}} \vee R_{\rho,\sigma} \\ &= E_{x_1, x_2\rho^\beta} \vee R_{\rho,\sigma} \end{aligned}$$

Furthermore the row period of $f|_{\rho_{i,j_2}\times\mathbf{n}}$ is equal to $\lambda_f(i,*)$. Therefore row b is equal to row $b\rho^{z\lambda_f(i,*)}$ for all $z\in\mathbb{Z}$. Let $k\in\lambda_f(i,*)$ be such that $\beta = k + z\,\lambda_f(i,*)$ for some $z\in\mathbb{Z}$. Then

$$\begin{aligned} E_{x_1,x_2\rho^\beta}\vee R_{\rho,\sigma} &= E_{x_1,x_2\rho^{k+z\,\lambda_f(i,*)}}\vee R_{\rho,\sigma}\\ &= E_{x_1,x_2\rho^k}\vee R_{\rho,\sigma}\end{aligned}$$

as required. The reverse implication follows immediately from the definition of $\{0,1\}^{E_{\rho_{i,j_1},\rho_{i,j_2}}}$ which is equal to a union of sets including $\{0,1\}^{E_{x_1,x_2'}\vee R_{\rho,\sigma}}$.

It remains to show that k is unique. To do this, let $0<k_1\leq k_2\leq\lambda_f(i,*)$ and write $x_2' = x_2\rho_1^k$ and $x_2'' = x_2\rho_2^k$. Now suppose that $f\in\{0,1\}^{E_{x_1,x_2'}\vee R_{\rho,\sigma}}$ and $f\in\{0,1\}^{E_{x_1,x_2''}\vee R_{\rho,\sigma}}$. Then we have that row x_1 is equal to row x'_2 and row x_2''. Since $x_2'' = x_2'\rho^{k_2-k_1}$ we have that row $x_1\rho^{k_2-k_1}$ is equal to row x_1. This is fine if $k_2-k_1=0$ but otherwise this implies the row period of $f|_{\mathrm{dom}(\rho_{i,j_1})}$ equal to $k_2-k_1<\lambda_f(i,*)$ which cannot be true the row period equals $\lambda_f(i,*)$. Therefore $k_1=k_2$. This shows the uniqueness of k, as required. □

Let $A\subseteq\mathscr{A}(\rho,\sigma)$ and let $\lambda\in\Lambda(A)$. With the help of Lemma 4.4.10 and Lemma 4.4.11 we can create a necessary and sufficient condition for a matrix $f\in\{0,1\}^{R_{\rho,\sigma}}$ to be in $\cap A$ with $\lambda_f=\lambda$ also being true. This condition will lead to a way to enumerate these matrices.

Lemma 4.4.12. *Let $(\rho,\sigma)\in S_m\times S_n$. Let $f\in\{0,1\}^{R_{\rho,\sigma}}$. Let $A\subseteq\mathscr{A}(\rho,\sigma)$. For all $i\in\mu(\mathbf{A})$ and $1<j\leq r_i(A)$ let $x_{i,j}$ be some element of $\mathrm{dom}(\rho_{i,j})$. For all $i\in\nu(\mathbf{A})$ and $1<j\leq s_i(A)$ let $y_{i,j}$ be some element of $\mathrm{dom}(\sigma_{i,j})$. Let $\lambda\in\Lambda(A)$. Then $f\in\cap A$ and $\lambda_f=\lambda$ if and only if the following conditions all hold:*

(i) *For each $(i,j)\in\mu(\mathbf{A})\times\nu(\mathbf{A})$ we have that $f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})}\in\zeta(\rho,\sigma,\rho_{i,1},\sigma_{j,1},\lambda(i,j))$.*

(ii) *For each $i\in\mu(\mathbf{A})$ and $1<j\leq r_i(A)$ there exists $\alpha_{i,j}\in\lambda(\mathbf{i},*)$ such that*

$$f\in\{0,1\}^{E_{x_{i,1},x'_{i,j}}}$$

where $x'_{i,j} = x_{i,j}\rho^{\alpha_{i,j}}$.

(iii) *For each $i\in\nu(\mathbf{A})$ and $1<j\leq s_i(A)$ there exists $\beta_{i,j}\in\lambda(*,\mathbf{i})$ such that*

$$f\in\{0,1\}^{F_{y_{i,1},y'_{i,j}}}$$

where $y'_{i,j} = y_{i,j}\sigma^{\beta_{i,j}}$.

Proof. Recall that $\cap A=\cap\bar{A}$. Therefore $f\in\cap A$ if and only if $f\in\cap\bar{A}$. Lemma 4.4.11 tells us that (ii) is equivalent to the statement: for each $i\in\mu(\mathbf{A})$ and $1<j\leq r_i(A)$ we have that

$f \in \{0,1\}^{E_{\rho_{i,1},\rho_{i,j}}}$. Similarly, the analogue to Lemma 4.4.11 for columns tells us that (iii) is equivalent to the statement: for each $i \in \nu(\mathbf{A})$ and $1 < j \leq s_i(A)$ we have that $f \in \{0,1\}^{F_{\sigma_{i,1},\sigma_{i,j}}}$.

Let us begin by assuming $f \in \cap A$ and $\lambda_f = \lambda$ to prove the forward implication. Consider the alternative statement for (ii) attained by using Lemma 4.4.11. We certainly have that $\{0,1\}^{E_{\rho_{i,1},\rho_{i,j}}} \in \bar{A}$ because $u_{i,1}$ and $u_{i,j}$ are in the same component of $G_R(A)$. Thus $f \in \{0,1\}^{E_{\rho_{i,1},\rho_{i,j}}}$ as required. By an analogous argument (iii) is also implied by our assumptions. To show (i) holds recall that, by definition, $\zeta(\rho,\sigma,\rho_{i,1},\sigma_{j,1},\lambda(i,j))$ is the collection

$$\{f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})} : f \in \{0,1\}^{R_{\rho,\sigma}} \text{ and the row period of } f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})} \text{ is } \lambda(i,j)\}.$$

Well since $\lambda_f = \lambda$ the row period of $f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})}$ is certainly $\lambda(i,j)$ and $f \in \{0,1\}^{R_{\rho,\sigma}}$ follows from $f \in \cap A$.

To prove the reverse implication, assume all of (i), (ii), and (iii) hold. First note that Lemma 4.4.10 tells us that (i) implies $\lambda_f = \lambda$. To complete the proof, we will show that $f \in B \subseteq \mathscr{A}(\rho,\sigma)$ such that $\bar{A} \subseteq \bar{B}$, which will show that $f \in \cap\bar{B} \subseteq \cap A$. By Lemma 4.4.11 we have that for all $i \in \mu(\mathbf{A})$ and $1 < j < r_i(A)$ we have $f \in \{0,1\}^{E_{\rho_{i,1},\rho_{i,j}}}$. Let

$$B_1 = \{\{0,1\}^{E_{\rho_{i,1},\rho_{i,j}}} : i \in \mu(\mathbf{A}) \text{ and } 1 < j \leq r_i(A)\}.$$

Note that $f \in \cap B_1$. Then

$$\bar{B}_1 = \{\{0,1\}^{E_{\rho_{i,j_1},\rho_{i,j_2}}} : i \in \mu(\mathbf{A}) \text{ and } 1 \leq j_1 < j_2 \leq r_i(A)\}$$

and we also have that $f \in \cap\bar{B}_1$. Similarly, let

$$B_2 = \{\{0,1\}^{F_{\sigma_{i,1},\sigma_{i,j}}} : i \in \nu(\mathbf{A}) \text{ and } 1 < j \leq s_i(A)\}$$

and notice that $f \in \cap B_1$. Then

$$\bar{B}_2 = \{\{0,1\}^{F_{\sigma_{i,j_1},\sigma_{i,j_2}}} : i \in \nu(\mathbf{A}) \text{ and } 1 \leq j_1 < j_2 \leq s_i(A)\}$$

and we also have that $f \in \cap\bar{B}_2$.

Lemma 4.4.10 shows that (i) implies $f \in \{0,1\}^{E_{\rho_{i,j},\lambda(i,*)}}$ for all $i \in \mu(\mathbf{A})$ and $j \in \mathbf{r_i}(\mathbf{A})$. Lemma 4.3.20 states that $f \in \{0,1\}^{E_{\rho_i,k}}$ if and only if the row period of $f|_{\mathrm{dom}(\rho_i)\times\mathbf{n}}$ divides k. It follows that if k_1 divides k_2, then $\{0,1\}^{E_{\rho_i,k_1}}$ is a subset of $\{0,1\}^{E_{\rho_i,k_2}}$. It is certainly true that $\lambda(i,*)$ divides $\delta_{R,i}(A)$ and so $f \in \{0,1\}^{E_{\rho_{i,j},\delta_{R,i}(A)}}$. A similar argument shows that

$f \in \{0,1\}^{F_{\sigma_{i,j},\delta_{C,i}(A)}}$ for all $i \in \nu(\mathbf{A})$. Now let

$$B_3 = \{\{0,1\}^{E_{\rho_{i,j},\delta_{R,i}(A)}} : i \in \mu(\mathbf{A}) \text{ and } j \in \mathbf{r_i}(\mathbf{A})\}.$$

We certainly have $f \in \cap B_3$. Then

$$\bar{B}_3 = \{\{0,1\}^{E_{\rho_{i,j},k}} : i \in \mu(\mathbf{A}), j \in \mathbf{r_i}(\mathbf{A}), \delta_{R,i}(A) \text{ divides } k, k \text{ divides } |\rho_{i,j}|, \text{ and } k \neq |\rho_{i,j}|\}\}$$

and $f \in \cap \bar{B}_3$ follows. Similarly, let

$$B_4 = \cup\{\{0,1\}^{F_{\sigma_{i,j},\delta_{C,i}(A)}} : i \in \nu(\mathbf{A}) \text{ and } j \in \mathbf{s_i}(\mathbf{A})\}.$$

Then

$$\bar{B}_4 = \{\{0,1\}^{F_{\sigma_{i,j},k}} : i \in \nu(\mathbf{A}), j \in \mathbf{s_i}(\mathbf{A}), \delta_{C,i}(A) \text{ divides } k, k \text{ divides } |\sigma_{i,j}|, \text{ and } k \neq |\sigma_{i,j}|\}\}$$

and $f \in \cap \bar{B}_4$ follows.

Now set $B = \bar{B}_1 \cup \bar{B}_2 \cup \bar{B}_3 \cup \bar{B}_4$. Since $f \in \cap \bar{B}_i$ for each $1 \leq i \leq 4$ we have that $f \in \cap B$. We also have that $\bar{B} = B$ and $B = \bar{A}$. Therefore (i), (ii) and (iii) imply that $f \in \cap A = \cap \bar{A} = \cap B$, as required. □

Given $\lambda \in \Lambda(A)$, we now show that Lemma 4.4.12 gives us a unique description of each matrix in $\{f \in \cap A : \lambda_f = \lambda\}$ in terms of the constants $\{\alpha_{i,j} : i \in \mu(\mathbf{A}), j \in \mathbf{r_i}\}$, $\{\beta_{i,j} : i \in \nu(\mathbf{A}), j \in \mathbf{s_i}\}$, and the sub-matrices on the domains $\text{dom}(\rho_{i,1}) \times \text{dom}(\sigma_{j,1})$. Furthermore we show that any description of the type described corresponds to a matrix in $\{f \in \cap A : \lambda_f = \lambda\}$. In other words, there is a one to one correspondence between $\{f \in \cap A : \lambda_f = \lambda\}$ and the combinations of choices for all the different constants $\alpha_{i,j}$, $\beta_{i,j}$, together with the sub-matrices on the domains $\text{dom}(\rho_{i,1}) \times \text{dom}(\sigma_{j,1})$. This will allow us to enumerate the former by enumerating the later.

Lemma 4.4.13. *Let $(\rho, \sigma) \in S_m \times S_n$. Let $A \subseteq \mathscr{A}(\rho, \sigma)$. For all $i \in \mu(\mathbf{A})$ and $1 < j \leq r_i(A)$ let $x_{i,j}$ be some element of $\text{dom}(\rho_{i,j})$. For all $i \in \nu(\mathbf{A})$ and $1 < j \leq s_i(A)$ let $y_{i,j}$ be some element of $\text{dom}(\sigma_{i,j})$. Let $\lambda \in \Lambda(A)$. Let $f, g \in \cap A$ be such that $\lambda_f = \lambda_g = \lambda$. Then $f = g$ if and only if the following conditions all hold:*

(i) *Let $i \in \mu(\mathbf{A})$ and $1 < j \leq r_i(A)$. Assume $a, b \in \lambda(\mathbf{i}, *)$ satisfy*

$$f \in \{0,1\}^{E_{x_{i,1},x'}} \text{ and } g \in \{0,1\}^{E_{x_{i,1},x''}}$$

where $x' = x_{i,j}\rho^a$ and $x'' = x_{i,j}\rho^b$. Then $a = b$.

(ii) *Let* $i \in \nu(\mathbf{A})$ *and* $1 < j \leq s_i(A)$. *Assume* $a, b \in \lambda(*, \mathbf{i})$ *satisfy*

$$f \in \{0,1\}^{F_{y_{i,1},y'}} \text{ and } g \in \{0,1\}^{F_{y_{i,1},y''}}$$

where $y' = y_{i,j}\sigma^a$ *and* $y'' = y_{i,j}\sigma^b$. *Then* $a = b$.

(iii) *Let* $(i,j) \in \mu(\mathbf{A}) \times \nu(\mathbf{A})$. *Then* $f|_{\mathrm{dom}(\rho_{i,1}) \times \mathrm{dom}(\sigma_{j,1})}$ *equals* $g|_{\mathrm{dom}(\rho_{i,1}) \times \mathrm{dom}(\sigma_{j,1})}$.

Proof. We will begin with the forward implication. Take $f \in \cap A$ such that $\lambda_f = \lambda$. If we have $a, b \in \lambda(\mathbf{i}, *)$ such that

$$f \in \{0,1\}^{E_{x_{i,1},x'}} \text{ and } f \in \{0,1\}^{E_{x_{i,1},x''}}$$

as described in (i), then $a = b$ by the uniqueness condition of Lemma 4.4.11. Analogously, if we have $a, b \in \lambda(*, \mathbf{i})$ as described in (ii), then $a = b$. Finally (iii) is clear because $f = g$ so a sub-matrix of f and a sub-matrix of g over the same domain will also be equal.

Now we prove the reverse implication. We choose to prove the contrapositive. That means we will show that if $f, g \in \cap A$ such that $\lambda_f = \lambda_g = \lambda$ and (i), (ii), and (iii) do not all hold, then $f \neq g$. Therefore assume that $f, g \in \cap A$ such that $\lambda_f = \lambda_g = \lambda$ and that (i), (ii), and (iii) do not all hold. Then either (i), (ii), or (iii) does not hold. Assume (i) does not hold and let $i \in \mu(\mathbf{A})$ and $1 < j \leq r_i(A)$ be such that there exists $a, b \in \lambda(\mathbf{i}, *)$ so that (i) does not hold. Then we have

$$f \in \{0,1\}^{E_{x_{i,1},x'}} \text{ and } g \in \{0,1\}^{E_{x_{i,1},x''}}$$

where $x' = x_{i,j}\rho^a$, $x'' = x_{i,j}\rho^b$, and $a \neq b$. Without loss of generality, assume that $a > b$. If row $x_{i,1}$ of f does not equal row $x_{i,1}$ of g we are done. Otherwise, assume they are equal. Then row $x_{i,1}$ of f equals row x' of f and row $x_{i,1}$ of g equals row x'' of g. We have that $x'' = x'^{a-b}$ and so row x' of g cannot equal row x'' of g since $0 < a - b < \lambda(i, *)$. In particular, $a - b$ is not a multiple of the row period of $g|_{\mathrm{dom}(\rho_{i,j}) \times \mathbf{n}}$. Therefore not (i) implies $f \neq g$. By an analogous argument, not (ii) implies $f \neq g$. Finally, not (iii) clearly implies $f \neq g$. This completes the proof □

Now we have an alternative representation of the elements of the set $\{f \in \cap A : \lambda_f = \lambda\}$. Thus we can enumerate this set by enumerating the possibilities for: the constants $\{\alpha_{i,j} : i \in \mu(\mathbf{A}), j \in \mathbf{r_i}\}$, the constants $\{\beta_{i,j} : i \in \nu(\mathbf{A}), j \in \mathbf{s_i}\}$, and the sub-matrices on the domains $\mathrm{dom}(\rho_{i,1}) \times \mathrm{dom}(\sigma_{j,1})$ which a matrix in $\{f \in \cap A : \lambda_f = \lambda\}$ may have.

Lemma 4.4.14. *Let* $(\rho, \sigma) \in S_m \times S_n$. *Let* $A \subseteq \mathscr{A}(\rho, \sigma)$ *and let* $\lambda \in \Lambda(A)$. *Then the number of* $f \in \cap A$ *such that* $\lambda_f = \lambda$ *is*

$$\left(\prod_{(i,j)\in\mu(\mathbf{A})\times\nu(\mathbf{A})}\omega(\lambda(i,j))\right)\cdot\left(\prod_{i\in\mu(\mathbf{A})}\lambda(i,*)^{r_i(A)-1}\right)\cdot\left(\prod_{j\in\nu(\mathbf{A})}\lambda(*,j)^{s_j(A)-1}\right)$$

Proof. Lemma 4.4.12 shows us that each matrix $f\in\cap A$ such that $\lambda_f=\lambda$ can be characterised by the sub-matrices on the domains $\{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1}) : i\in\mu(\mathbf{A}) \text{ and } j\in\nu(\mathbf{A})\}$; together with a selection of constants $\alpha_{i,j}\in\lambda(\mathbf{i},*)$ for $i\in\mu(\mathbf{A})$ and $1<j\leq r_i(A)$, and $\beta i,j\in\lambda(*,\mathbf{i})$ for $i\in\nu(\mathbf{A})$ and $1<j\leq s_i(A)$. Furthermore Lemma 4.4.13 shows there is a one to one correspondence between these selections and the matrices f in $\cap A$ with $\lambda_f=\lambda$. Thus we can enumerate the latter by enumerating the former. There are $\lambda(i,*)$ choices for $\alpha_{i,j}$, therefore there are

$$\prod_{i\in\mu(\mathbf{A})}\lambda(i,*)^{r_i(A)-1}$$

ways to choose all the $\alpha_{i,j}$ for $i\in\mu(\mathbf{A})$ and $1<j\leq r_i(A)$. Similarly, there are

$$\prod_{i\in\mu(\mathbf{A})}\lambda(i,*)^{r_i(A)-1}$$

ways to choose all the $\beta i,j$ for $i\in\nu(\mathbf{A})$ and $1<j\leq s_i(A)$. Finally, there are $\omega(\lambda(i,j))=|\zeta(\rho,\sigma,\rho_{i,1},\sigma_{j,1},\lambda(i,j))|$ ways to choose an element of $\zeta(\rho,\sigma,\rho_{i,1},\sigma_{j,1},\lambda(i,j))$. Therefore there are

$$\prod_{(i,j)\in\mu(\mathbf{A})\times\nu(\mathbf{A})}\omega(\lambda(i,j))$$

ways to choose all the sub-matrices on the domains $\{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1}) : i\in\mu(\mathbf{A}) \text{ and } j\in\nu(\mathbf{A})\}$. The result follows. □

For any $A\in\mathscr{A}(\rho,\sigma)$, and $\lambda\in\Lambda(A)$ we can now enumerate the collection $\{f\in\cap A : \lambda_f=\lambda\}$. We will define

$$\Omega(A,\lambda)=\left(\prod_{(i,j)\in\mu(\mathbf{A})\times\nu(\mathbf{A})}\omega(\lambda(i,j))\right)\cdot\left(\prod_{i\in\mu(\mathbf{A})}\lambda(i,*)^{r_i(A)-1}\right)\cdot\left(\prod_{j\in\nu(\mathbf{A})}\lambda(*,j)^{s_j(A)-1}\right) \quad (4.5)$$

for convenience. Now to enumerate $\cap A$ we simply sum the $Omega(A,\lambda)$ over all $\lambda\in\Lambda(A)$.

Theorem 4.4.15. *Let $(\rho,\sigma)\in S_m\times S_n$ and A be a subset of $\mathscr{A}(\rho,\sigma)$. Then*

$$\left|\bigcap_{a\in A}a\right|=\sum_{\lambda\in\Lambda(A)}\Omega(A,\lambda),$$

where Ω maps $A\in\mathscr{A}(\rho,\sigma)$ and $\lambda\in\Lambda(A)$ to the size of $\{f\in\cap A : \lambda_f=\lambda\}$.

Proof. The sets $\{\{f \in \cap A : \lambda_f = \lambda\} : \lambda \in \Lambda(A)\}$ are and their union is equal to $\cap A$. The result follows since $\Omega(A, \lambda)$ is equal to the size of $\{f \in \cap A : \lambda_f = \lambda\}$. □

To summarise, in this section we have shown how to enumerate $\cap A$ by enumerating the classes of the partition Q_A where elements f, g of $\cap A$ are in the same class if and only if $\lambda_f = \lambda_g$. We enumerated each of these classes by a combinatorial method. Using this result we are halfway to being able to apply the inclusion-exclusion principle with terms collected via the partition $P(\rho, \sigma)$ as shown in Equation 4.4. Now, given $p \in P(\rho, \sigma)$ we can take some representative A_p of p and determine $\left|\cap A_p\right|$. In the next section we will show how to determine the coefficient k_p.

4.4.3 Counting graphs by edge and label parity

The focus of this section will be to take p from $P(\rho, \sigma)$, the partition of $\mathscr{P}(\mathscr{A}(\rho, \sigma))$ which we determined earlier such that elements of the same class have equal sized intersection, and determine the coefficient k_p from Equation 4.4. Recall that

$$k_p = \sum_{A \in p} (-1)^{|A|+1}.$$

That is to say, k_p is the sum over all $A \in p$ such that we add 1 if A has even cardinality and -1 if A has odd cardinality. The property of an integer being odd or even is sometimes called its *parity*. We will refer to odd integers as having parity -1 and even integers as having parity 1. Note that the parity of an integer z is equal to $(-1)^z$. We will define the *parity* of a finite set X to be $(-1)^{|X|}$. With that definition, k_p is equal to the sum of the parities of all A in p.

Recall that $A \subseteq \mathscr{A}(\rho, \sigma)$ has an associated pair of graphs $(G_R(A), G_C(A))$. Each element of A corresponds to either an edge or a vertex label of one of $G_R(A)$ or $G_C(A)$. Thus the cardinality of A equals the sum of the following four counts: the number of edges in $G_R(A)$, plus the number of edges in $G_R(A)$, plus the sum of the sizes of all vertex labels in $G_R(A)$, plus the sum of the sizes of all vertex labels in $G_C(A)$. Therefore the parity of the set A is the product of the parities of all these counts.

Remember that the class of $A \subseteq \mathscr{A}(\rho, \sigma)$ in the partition $P(\rho, \sigma)$ is

$$[A]_{P(\rho,\sigma)} = \{B \subseteq \mathscr{A}(\rho, \sigma) : \exists (\mu, \nu) \in C_{S_m \times S_n}((\rho, \sigma)) : \bar{A} = \bar{B} \cdot (\mu, \nu)\}.$$

Consider the set

$$\{\bar{B} : B \in [A]_{P(\rho,\sigma)}\}$$

and denote its size by $\gamma(A)$ and its elements by $A_1,\ldots,A_{\gamma(A)}$. Then we have

$$[A]_{P(\rho,\sigma)} = \bigcup_{i=1}^{\gamma(A)} \{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{A}_i = \bar{B}\}$$

and furthermore this is a disjoint union, i.e.

$$\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{A}_i = \bar{B}\} \cap \{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{A}_j = \bar{B}\} = \emptyset$$

for all $1 \leq i < j \leq \gamma(A)$. In this section will show how to determine the sum of the parities of the elements in a set of the form $\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{A}_i = \bar{B}\}$. The sum of the parities of the elements in the set $[A]_{P(\rho,\sigma)}$ is equal to $\gamma(A)$ times the sum for one of those sets. Therefore our approach will be to determine the sum of the parities of the elements in $\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}$ and multiply that by $\gamma(A)$. We show how to determine $\gamma(A)$ now.

The set $\{\bar{B} \subseteq \mathscr{A}(\rho,\sigma) : B \in [A]_{P(\rho,\sigma)}\}$ is equal to the orbit of $\bar{A}$ with respect to the action of $C_{S_m \times S_n}((\rho,\sigma))$. This follows from the definition of $P(\rho,\sigma)$. Thus $\gamma(A)$ is equal to the cardinality of this orbit. To have intuition about this action consider the corresponding action on the graph pair $(G_R(\bar{A}), G_C(\bar{A}))$. This action can swap any two vertices in one of the graphs corresponding to cycles of the same length. Not all swaps will produce a distinct graph. The stabilizer of this action will contain any swap of two vertices in the same connected component, as well as any swap of two vertices in identical components (that is, components of the same size with the same greatest common divisors of labels). To determine the size of the orbit we will apply the orbit-stabilizer theorem.

Lemma 4.4.16. *Let $(\rho,\sigma) \in S_m \times S_n$ and let $A \subseteq \mathscr{A}(\rho,\sigma)$. Then*

$$\gamma(A) = \frac{|C_{S_m}(\rho)| \cdot |C_{S_n}(\sigma)|}{|Stab_{C_{S_m}((\rho))}(G_R(\bar{A}))| \cdot |Stab_{C_{S_n}((\sigma))}(G_C(\bar{A}))|}.$$

Proof. Since $\gamma(A)$ is equal to the size of the orbit of $\bar{A}$ with respect to $C_{S_m \times S_n}((\rho,\sigma))$ we apply the orbit-stabilizer theorem to obtain

$$\gamma(A) = \frac{|C_{S_m \times S_n}((\rho,\sigma))|}{|Stab_{C_{S_m \times S_n}((\rho,\sigma))}((G_R(\bar{A}), G_C(\bar{A})))|}.$$

The actions of the two factors S_m and S_n are independent of each other. Therefore we have

$$|C_{S_m \times S_n}((\rho,\sigma))| = |C_{S_m}(\rho)| \cdot |C_{S_n}(\sigma)|$$

and

$$|Stab_{C_{S_m\times S_n}((\rho,\sigma))}((G_R(\bar{A}),G_C(\bar{A})))| = |Stab_{C_{S_m}((\rho))}(G_R(\bar{A}))|\cdot|Stab_{C_{S_n}((\sigma))}(G_C(\bar{A}))|.$$

This completes the proof. □

Let us denote

$$\gamma(A,\rho) = \frac{|C_{S_m}(\rho)|}{|Stab_{C_{S_m}((\rho))}(G_R(\bar{A}))|},$$

and

$$\gamma(A,\sigma) = \frac{|C_{S_n}(\sigma)|}{|Stab_{C_{S_n}((\sigma))}(G_C(\bar{A}))|}.$$

Then we have that $\gamma(A) = \gamma(A,\rho)\gamma(A,\sigma)$ and we will be able to calculate $\gamma(A,\rho)$ and $\gamma(A,\sigma)$ by analogous methods. The following lemma shows how best to enumerate these expressions.

Lemma 4.4.17. *Let $(\rho,\sigma) \in S_m \times S_n$ and let $A \subseteq \mathscr{A}(\rho,\sigma)$. Then*

$$\gamma(A,\rho) = \frac{\prod_{i=1}^{m} a_i!}{\prod_{i=1}^{m}\prod_{j=1}^{a_i}(j!)^{k_{i,j}}\prod_{z=1}^{m} k_{i,j,z}!}.$$

where a_i is the number of i-cycles of ρ; $k_{i,j}$ is the number of connected components of $G_R(A)$ of size j containing vertices corresponding to cycles of ρ of length i; and $k_{i,j,z}$ is the number of connected components of $G_R(A)$ of size j containing vertices corresponding to cycles of ρ of length i with label greatest common divisor equal to z.

Proof. Let a_i denote the number of i-cycles of ρ. Recall that $\rho = \rho_1\cdots\rho_r$ where $\rho_1,\ldots,\rho_r$ are the cycles (including those of length one) which compose ρ. The permutations $\{\prod_{i=1}^{r}\rho_i^{k_i} : k_1 \in |\rho_{\mathbf{1}}|,\ldots,k_r \in |\rho_{\mathbf{r}}|\}$ are all elements of $C_{S_m}(\rho)$. There are $\prod_{i=1}^{m} i^{a_i}$ permutations of this type. Furthermore, $C_{S_m}(\rho)$ can act by permuting cycles of the same length and all such permutations are possible. There are $a_i!$ ways to permute the i-cycles of ρ for each $i \in \mathbf{m}$ and thus $\prod_{i=1}^{m} a_i!$ ways to permute cycles of ρ. The elements of the centraliser $C_{S_m}(\rho)$ can all be expressed as $(\prod_{i=1}^{r}\rho_i^{k_i})f$ where f is a permutation of the cycles of ρ of the type just discussed. Thus $C_{S_m}(\rho)$ has size $\prod_{i=1}^{m} i^{a_i}a_i!$.

Now we determine the subset of $C_{S_m}(\rho)$ which stabilizes $G_R(\bar{A})$. First, note that all permutations in $\{\prod_{i=1}^{r}\rho_i^{k_i} : k_1 \in |\rho_{\mathbf{1}}|,\ldots,k_r \in |\rho_{\mathbf{r}}|\}$ stabilize $G_R(\bar{A})$ because they do nothing - they permute no cycles of ρ and thus permute none of the vertices of $G_R(\bar{A})$. Next note that by the definition of $\bar{A}$ each component of $G_R(\bar{A})$ is a complete graph with all vertices having the same label. Thus if we swap two vertices in a component the graph is unchanged. In fact any permutation of the cycles of ρ which corresponds to permuting vertices of $G_R(\bar{A})$ in a way that

leaves the connected components are unchanged must fix $G_R(\bar{A})$. There are $\prod_{i=1}^{m}\prod_{j=1}^{a_i}(j!)^{k_{i,j}}$ such permutations of the vertices since there are $j!$ ways to permute the vertices of a connected component of size j. Finally note that if two connected components of equal size contain vertices corresponding to cycles of the same length and their vertex labels have the same greatest common divisor then swapping these connected components fixes $G_R(\bar{A})$. If there are $k_{i,j,z}$ many connected components of size j containing vertices corresponding to cycles of length i and with label greatest common divisor z then there are $k_{i,j,z}!$ ways to permute these connected components. Combining all the ways to fix $G_R(\bar{A})$ discussed we determine:

$$Stab_{C_{S_m}(\rho)}(G_R(\bar{A}) = \prod_{i=1}^{m} i^{a_i} \prod_{j=1}^{a_i} (j!)^{k_{i,j}} \prod_{z=1}^{m} k_{i,j,z}!.$$

Finally we notice that $\prod_{i=1}^{m} i^{a_i}$ divides both $|C_{S_m}(\rho)|$ and $|Stab_{C_{S_m}((\rho))}(G_R(\bar{A}))|$ so we remove it from the numerator and denominator of

$$\frac{|C_{S_m}(\rho)|}{|Stab_{C_{S_m}((\rho))}(G_R(\bar{A}))|} = \frac{\prod_{i=1}^{m} i^{a_i} a_i!}{\prod_{i=1}^{m} i^{a_i} \prod_{j=1}^{a_i} (j!)^{k_{i,j}} \prod_{z=1}^{m} k_{i,j,z}!}$$

to obtain the result. □

Now we can move on to determining the sum of the parities of the elements in a set of the form $\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{A}_i = \bar{B}\}$. We begin by considering just the subset of an element $A \subset \mathscr{A}(\rho,\sigma)$ which corresponds to the edges of the graphs in the corresponding graph pair $G(A) = (G_R(A), G_C(A))$. For any graph G we will let $E(G)$ denote the number of edges of G. The *edge parity* of a graph G will refer to parity of the set of edges, i.e. the number $(-1)^{E(G)}$. When we refer to the edge parity of the pair graphs $(G_R(A), G_C(A))$ for some $A \subseteq \mathscr{A}(\rho,\sigma)$ we will be referring to the product of the edge parities of the two graphs. Lemma 4.4.18 allows us to calculate the edge parity of a collection of graphs in a specific case which is an important piece of the general case.

Lemma 4.4.18 ([22]). *Let C_n denote the connected graphs on n vertices. Then the following equation counts these graphs by parity of number of edges*

$$\sum_{G \in C_n} (-1)^{E(G)} = (-1)^{n-1}(n-1)!.$$

Herein we will denote $\sum_{G \in C_n} (-1)^{E(G)}$ by $\psi(n)$. To see the application of Lemma 4.4.18, consider $A \subseteq \mathscr{A}(\rho,\sigma)$ such that the corresponding pair of graphs $(G_R(A), G_C(A))$ have no

vertex labels and only one connected component of size greater than one. Say this component is in $G_R(A)$ and is the component $K_{R,i}(A)$. Recall that we denote the number of vertices of $K_{R_i}(A)$ and $K_{C,j}(A)$ by $r_i(A)$ and $s_j(A)$, respectively. Then A is a subset of $\{\{0,1\}^{E_{\rho_{i,j}},E_{\rho_{i,k}}} : 1 \leq j < k \leq r_i(A)\}$. If we consider the collection $\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}$ then the sum by edge parity of the collection of graph pairs corresponding to this set is $\psi(r_i(A))$. We may also apply Lemma 4.4.18 to the case where the subset A of $\mathscr{A}(\rho,\sigma)$ is such that the corresponding pair of graphs $(G_R(A), G_C(A))$ has any edge set but has no vertex labels.

Lemma 4.4.19. *Let A of $\mathscr{A}(\rho,\sigma)$ be such that $G_R(A)$ and $G_C(A)$ have no vertex labels. Let $r_i(A)$ denote the number of vertices of the connected component $K_{R,i}(A)$ of $G_R(A)$. Let $s_i(A)$ denote the number of vertices of the connected component $K_{C,i}(A)$ of $G_C(A)$. Then the sum by edge parity of $\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}$ is equal to*

$$\left[\prod_{i=1}^{\mu(A)} \psi(r_i(A))\right] \cdot \left[\prod_{i=1}^{\nu(A)} \psi(s_i(A))\right].$$

Proof. Let us write $Y = \{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}$. All elements of Y have the same connected components but varying edge sets. Consider the sets $\{E(K_{R,i}(B)) : B \in Y\}$ for $1 \leq i \leq \mu(A)$ which contain all the edge sets of the connected components $K_{R,i}(B)$ of a graph pair corresponding to $B \in Y$. Similarly, consider the analogous sets $\{E(K_{C,i}(B)) : B \in Y\}$ for $1 \leq i \leq \nu(A)$. Clearly the edge set of one connected component is independent of the edge set of another amongst elements of Y. Therefore there are

$$\left[\prod_{i=1}^{\mu(A)} |\{E(K_{R,i}(B)) : B \in Y\}|\right] \cdot \left[\prod_{i=1}^{\nu(A)} |\{E(K_{C,i}(B)) : B \in Y\}|\right]$$

combined possibilities for the edge sets of $G_R(A)$ and $G_C(A)$. For some $B \in Y$ the edge parity of the graph pair $(G_R(B), G_C(B))$ is the product of the edge parities of all the components

$$\left[\prod_{i=1}^{\mu(B)} (-1)^{|E(K_{R,i}(B))|}\right] \cdot \left[\prod_{i=1}^{\nu(B)} (-1)^{|E(K_{C,i}(B))|}\right]$$

Then we can deduce

$$\begin{aligned}
&\sum_{A\in Y}\left(\left[\prod_{i=1}^{\mu(A)}(-1)^{|E(K_{R,i}(A))|}\right]\cdot\left[\prod_{i=1}^{\nu(A)}(-1)^{|E(K_{C,i}(A))|}\right]\right)\\
&=\left[\sum_{\{G_R(A):A\in Y\}}\prod_{i=1}^{\mu(A)}(-1)^{|E(K_{R,i}(A))|}\right]\cdot\left[\sum_{\{G_C(A):A\in Y\}}\prod_{i=1}^{\nu(A)}(-1)^{|E(K_{C,i}(A))|}\right]\\
&=\left[\sum_{\{E(K_{R,1}(A)):A\in Y\}}\cdots\sum_{\{E(K_{R,\mu(A)}(A)):A\in Y\}}\prod_{i=1}^{\mu(A)}(-1)^{|E(K_{R,i}(A))|}\right]\\
&\quad*\left[\sum_{\{E(K_{C,1}(A)):A\in Y\}}\cdots\sum_{\{E(K_{\{C,\nu(A)}(A)):A\in Y\}}\prod_{i=1}^{\nu(A)}(-1)^{|E(K_{C,i}(A))|}\right]\\
&=\left[\prod_{i=1}^{\mu(A)}\sum_{\{E(K_{R,i}(A)):A\in Y\}}(-1)^{|E(K_{R,i}(A)))|}\right]\cdot\left[\prod_{i=1}^{\nu(A)}\sum_{\{E(K_{C,i}(A)):A\in Y\}}(-1)^{|E(K_{C,i}(A))|}\right]\\
&=\left[\prod_{i=1}^{\mu(A)}\psi(r_i(A))\right]\cdot\left[\prod_{i=1}^{\nu(A)}\psi(s_i(A))\right]
\end{aligned}$$

The first and second equalities follow from the fact that the edge sets of each component are independent of each other. The third equality follows from the identity

$$\sum_{(x_1,\ldots,x_k)\in X_1\times\ldots\times X_k}\prod_{i=1}^{k}f(x_i)=\prod_{i=1}^{k}\sum_{x_i\in X_i}f(x_i)$$

where f is some function, in our case $f(x)=(-1)^{|x|}$. The fourth equality follows from Lemma 4.4.18 being applied to each connected component of $G_R(A)$ and $G_C(A)$ as, for example, the set $\{E(K_{R,i}(B)):B\in Y\}$ corresponds with the collection of all connected graphs on $r_i(A)$ vertices. □

Herein we will denote

$$\Psi(A)=\left[\prod_{i=1}^{\mu(A)}\psi(r_i(A))\right]\cdot\left[\prod_{i=1}^{\nu(A)}\psi(s_i(A))\right] \tag{4.6}$$

to allow us to refer to this expression more easily. Although we have ignored the cases where vertices are labelled thus far, the function Ψ will play a key part in the solution to the general case. We begin by considering the simple case of a connected graph.

We will say that a graph G with n vertices labelled by $L_1,\ldots,L_n$ has *label parity* equal to the product of the parities of the label sets:

$$\prod_{i=1}^{n}(-1)^{|L_i|}.$$

We will now show how to calculate the sum of labelling parities of a certain collection of labellings of a connected graph. Let G be a connected graph with n vertices. For $x \in \mathbb{N}$ define $\mathscr{G}(G,x)$ to be the set of all graphs with vertex and edge set equal to G where the vertices are labelled by subsets of the proper divisors $\mathrm{div}(x)$ of x. For k a proper positive divisor of x we define $\mathscr{G}(G,x,k)$ to be the subset of $\mathscr{G}(G,x)$ containing only the graphs where the greatest common divisor of the union of the labels equals k. Denote the sum by label parity of the graphs in $\mathscr{G}(G,x,k)$ by $\theta(G,x,k)$. Note that the greatest common divisor of the union of the labels is not defined in the case where all labels are empty. We set $\theta(G,x,x) = 1$ intending to refer to that case.

Lemma 4.4.20. *Let G be a graph with n vertices, let $x \in \mathbb{N}$, and let k be a proper positive divisor of x. Then*

$$\theta(G,x,k) = (-1)^z$$

if and only if k is equal to $x/(p_1 * \ldots * p_z)$ for some prime numbers $p_1,\ldots,p_z$ which are all distinct. Otherwise

$$\theta(G,x,k) = 0.$$

Proof. We will need to verify the following result to proceed:

$$\sum_{z\in \mathrm{div}(x):k|z} \theta(G,x,z) = 0$$

which is quickly done by noticing that this is really the sum by edge parity of all labellings by subsets of $\{z \in \mathrm{div}(x) : k|z\}$. If x has $X = |\{z \in \mathrm{div}(x) : k|z\}|$ proper divisors, then there are $\binom{nX}{i}$ labellings such that the sum of the sizes of the labels is i. Therefore the sum over all possible labellings is

$$\sum_{L_1,\ldots,L_n\subseteq\{z\in\mathrm{div}(x):k|z\}} (-1)^{\sum_{i=1}^{n}|L_i|} = \sum_{i=0}^{nX}\binom{nX}{i}\cdot(-1)^i = 0 = \sum_{z\in\{z\in\mathrm{div}(x):k|z\}} \theta(G,x,k).$$

Rearranging, we obtain:

$$\theta(G,x,k) = -\sum_{z\in\{z\in\mathrm{div}(x):k|z\&z\neq k\}} \theta(G,x,y).$$

Now we will perform a proof by strong induction. We have shown that when $z = 0$ (i.e. $x = k$) that the statement for all $y \leq z$ we have that if k is any integer such that $k = x/p_1 \cdots p_y$ for primes $p_1, \ldots, p_z$, not necessarily distinct, then

$$\theta(G,x,k) = (-1)^z$$

if $p_1, \ldots, p_z$ are all distinct; otherwise

$$\theta(G,x,k) = 0.$$

Now assume this statement holds for some $z > 0$ and let $k = x/p_1 \cdots p_{z+1}$ for some primes $p_1, \ldots, p_{z+1}$. Then

$$\theta(G,x,k) = -\sum_{z \in \{z \in \mathrm{div}(x) : k \mid z \,\&\, z \neq k\}} \theta(G,x,y).$$

If we let $q_1, \ldots, q_\alpha$ be the distinct primes in $\{p_1, \ldots, p_{z+1}\}$, then $\alpha \neq z+1$ implies

$$\theta(G,x,k) = -\sum_{(w_1,\ldots,w_\alpha) \in \{0,1\}^\alpha} \theta(G,x,x/q_1^{w_1} \cdots q_\alpha^{w_\alpha}) = 0$$

since all other terms were equal to zero. Otherwise $x/q_1 \cdots q_\alpha = \theta(G,x,k)$ and should not be part of the sum on the RHS. Therefore

$$\begin{aligned}
\theta(G,x,k) &= -\sum_{(w_1,\ldots,w_\alpha) \in \{0,1\}^\alpha \setminus \{(1,\ldots,1)\}} \theta(G,x,x/q_1^{w_1} \cdots q_\alpha^{w_\alpha}) \\
&= \sum_{i=0}^{\alpha-1} \binom{\alpha}{i} (-1)^i \\
&= (-1)^\alpha
\end{aligned}$$

since there are $\binom{\alpha}{i}$ subsets of $\{q_1, \ldots, q_\alpha\}$ of size i for each i and using the inductive assumption. Therefore the statement holds for all divisors k of x. □

Since $\theta(G,x,k)$ does not depend on G we will simply write $\theta(x,k)$ to refer to the value of $\theta(G,x,k)$. The idea is that the labelling of graphs in $\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}$ can be considered independently of the edge set. We will denote by $\Theta(A)$ the sum by label parity of all graph pairs with the same edge sets as $(G_R(A), G_C(A))$. Then for any B satisfying $\bar{B} = \bar{A}$ we have $\Theta(B) = \Theta(B)$. Furthermore in the collection of graph pairs corresponding to $\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}$ any edge set may be matched with any labelling. Therefore we

can show the sum by combined edge and label parity of $\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}$ will be $\Psi(A) * \Theta(A)$. We now show how to define Θ in terms of θ.

For $A \subseteq \mathscr{A}(\rho,\sigma)$ let $\mathscr{B}(A)$ denote the subset of $\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}$ such that $B \in \mathscr{B}(A)$ if and only if $G_R(B)$ has the same edge set as $G_R(A)$ and $G_C(B)$ has the same edge set as $G_C(A)$. To sum the labelling parities of all pairs of graphs corresponding to element of $\mathscr{B}(A)$ we apply Lemma 4.4.20 to each connected component of the graph pair and take the product of all the label parities.

Lemma 4.4.21. *Let $(\rho,\sigma) \in S_m \times S_n$ and let $A \subseteq \mathscr{A}(\rho,\sigma)$. Then the sum of the labelling parities of the graphs corresponding to elements of $\mathscr{B}(A)$ is*

$$\left(\prod_{x=1}^{\mu(A)} \theta(|\rho_{i,1}|, \delta_{R,i}(A))\right) \cdot \left(\prod_{x=1}^{\nu(A)} \theta(|\sigma_{i,1}|, \delta_{C,i}(A))\right).$$

Proof. This follows by a similar argument to Lemma 4.4.19. The labelling of each connected component of a graph pair is independent from the labelling of any other component. The label parity of some $B \in \mathscr{B}(A)$ is equal to the product of the label parities of the connected components of $G_R(A)$ and $G_C(A)$. We can take all possible labellings of elements of $\mathscr{B}(A)$ and sum their labelling parities. Recall that $L_B : V(G_R(A)) \cup V(G_C(A)) \to \mathbb{N}$ denotes the labelling of the vertices of the graph pair $(G_R(B), G_C(B))$. Let us write $L_{K_{R,i}(B)} : V(K_{R,i}(B)) \to \mathbb{N}$ to denote the labelling of some connected component by $K_{R,i}(B)$ of $G_R(B)$, and similarly $L_{K_{C,i}(B)}$ will denote the labelling for a connected component of $G_C(B)$. Then we deduce:

$$
\begin{aligned}
&\sum_{\{L_B : B \in \mathscr{B}(A)\}} \prod_{x \in V(G_R(B)) \cup V(G_C(B))} (-1)^{|L_B(x)|} \\
&= \left[\sum_{\{L_{K_{R,1}(B)} : B \in \mathscr{B}(A)\}} \cdots \sum_{\{L_{K_{R,\mu(A)}(B)} : B \in \mathscr{B}(A)\}} \prod_{i=1}^{\mu(B)} \prod_{x \in V(K_{R,i}(B))} (-1)^{|L_{K_{R,i}(A)}(x)|} \right] \\
&\quad * \left[\sum_{\{L_{K_{C,1}(B)} : B \in \mathscr{B}(A)\}} \cdots \sum_{\{L_{K_{C,\nu(A)}(B)} : B \in \mathscr{B}(A)\}} \prod_{i=1}^{\nu(B)} \prod_{x \in V(K_{C,i}(B))} (-1)^{|L_{K_{C,i}(A)}(x)|} \right] \\
&= \left[\prod_{i=1}^{\mu(A)} \sum_{\{L_{K_{R,i}(B)} : B \in \mathscr{B}(A)\}} \prod_{x \in V(K_{R,i}(B))} (-1)^{|L_{K_{R,i}(A)}(x)|} \right] \\
&\quad * \left[\prod_{i=1}^{\nu(A)} \sum_{\{L_{K_{C,i}(B)} : B \in \mathscr{B}(A)\}} \prod_{x \in V(K_{C,i}(B))} (-1)^{|L_{K_{C,i}(A)}(x)|} \right] \\
&= \left[\prod_{i=1}^{\mu(A)} \theta(|\rho_{i,1}|, \delta_{R,i}(A)) \right] * \left[\prod_{i=1}^{\nu(A)} \theta(|\sigma_{i,1}|, \delta_{C,i}(A)) \right].
\end{aligned}
$$

The first equality follows from the fact that the labellings of each connected component are independent amongst elements of $\mathscr{B}(A)$, and an application of the identity

$$\sum_{(x_1,\ldots,x_k) \in X_1 \times \ldots \times X_k} \prod_{i=1}^{k} f(x_i) = \prod_{i=1}^{k} \sum_{x_i \in X_i} f(x_i)$$

where f is some function, in our case

$$f(K) = \prod_{z \in V(K)} (-1)^{|L_K(z)|}.$$

The second equality follows from another application of the aforementioned identity. The final equality follows by applying Lemma 4.4.20 to sum the labelling parities of all possible labellings of each connected component of $G_R(A)$ and $G_C(A)$. To see this, recall $\theta(x,k)$ equals the sum by labelling parity of all possible labellings of a connected graph with labels that are subsets of the proper divisors of x with greatest common divisor of labels equal to k. A labelling of every connected component of the graph pair is an combination of labellings of each individual component, and these component labellings are independent. For example, the set $\{L_{K_{R,i}(B)} : B \in \mathscr{B}(A)\}$ corresponds with the collection of all labellings of connected graphs

with $r_i(A)$ vertices which have labels that are subsets of $|\rho_i|$ and with greatest common divisor of labels equal to $\delta_{R,i}(A)$. □

For all $A \subseteq \mathscr{A}(\rho,\sigma)$ we will define

$$\Theta(A) = \left(\prod_{x=1}^{\mu(A)} \theta(|\rho_{i,1}|, \delta_{R,i}(A))\right) \cdot \left(\prod_{x=1}^{\nu(A)} \theta(|\sigma_{i,1}|, \delta_{C,i}(A))\right) \tag{4.7}$$

to allow us to more easily refer to this expression. Now we are ready to deduce the coefficients k_p for p in $P(\rho,\sigma)$ required for our application of the inclusion-exclusion principle.

Theorem 4.4.22. *Let $(\rho,\sigma) \in S_m \times S_n$ and let $A \subseteq \mathscr{A}(\rho,\sigma)$. Then the sum of the parities of the elements of the class $[A]_{P(\rho,\sigma)}$ of the partition $P(\rho,\sigma)$ is*

$$k_{[A]_{P(\rho,\sigma)}} = \gamma(A) \cdot \Psi(A) \cdot \Theta(A).$$

Proof. Let us denote $Y = \{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}$. First we claim that the edge sets and vertex labellings of elements of Y are independent. If $B_1, B_2 \in Y$, then there is $B_3 \in Y$ such that the edge sets are the same as B_1:

$$E(G_R(B_1)) = E(G_R(B_3)) \text{ and, } E(G_C(B_1)) = E(G_C(B_3))$$

and the vertex labels are the same as for B_2, i.e. $L_{B_2} = L_{B_3}$. Recall that $\mathscr{B}(B)$ denotes the subset of Y containing those C where the edge set of $G_R(B)$ is the same as the edge set of $G_R(C)$, and the edge set of $G_C(B)$ is the same as the edge set of $G_C(C)$.

Then we have that the sum by parities of the set Y is

$$\begin{aligned}
\sum_{B\in Y}(-1)^{|B|} &= \sum_{\{\mathscr{B}(B):B\in Y\}} \sum_{D\in\mathscr{B}(B)} (-1)^{|D|} \\
&= \sum_{\{\mathscr{B}(B):B\in Y\}} \sum_{D\in\mathscr{B}(B)} (-1)^{|E(G_R(D))|+|E(G_C(D))|} \cdot \prod_{v\in V(G_R(D))\cup V(G_C(D))} (-1)^{|L_D(v)|} \\
&= \sum_{\{\mathscr{B}(B):B\in Y\}} (-1)^{|E(G_R(B))|+|E(G_C(B))|} \sum_{D\in\mathscr{B}(B)} \prod_{v\in V(G_R(D))\cup V(G_C(D))} (-1)^{|L_D(v)|} \\
&= \sum_{\{\mathscr{B}(B):B\in Y\}} (-1)^{|E(G_R(B))|+|E(G_C(B))|} \sum_{D\in\mathscr{B}(B)} \Theta(A) \\
&= \Psi(A) \cdot \Theta(A).
\end{aligned}$$

Note that the first equality follows because Y is equal to the disjoint union of the set $\{\mathscr{B}(B) : B \in Y\}$, these sets are disjoint because each one contains all graphs pairs from Y with a certain

pair of edge sets. The second equality arises by substituting $|D|$ for the sum of the sizes of the two edge sets and all the vertex labels from the graph pair corresponding to D. The third equality follows because the edge sets are the same for all elements $D \in \mathscr{B}(B)$. The fourth equality applies Lemma 4.4.21 to count $\mathscr{B}(B)$ by label parity. The final equality follows by noticing the sets $\{\mathscr{B}(B) : B \in Y\}$ are in correspondence with the unique pairs of edge sets $\{(E(G_R(B)), E(G_C(B))) : B \in Y\}$ and taking the sum by edge parity of this collection using Lemma 4.4.19. We conclude the proof by recalling that $[A]_{P(\rho,\sigma)}$ is the disjoint union of

$$\{\{D \subseteq \mathscr{A}(\rho,\sigma) : \bar{D} = \bar{B}\}\} : B \in [A]_{P(\rho,\sigma)}.$$

Earlier in this section we showed that there are $\gamma(A)$ sets in this disjoint union and that each set has the same sum by parity. The result follows immediately. □

Theorem 4.4.22 together with Theorem 4.4.15 are all the tools we needed to apply Equation 4.4 to determine the size of the union of $\mathscr{A}(\rho,\sigma)$, which is equal to the number of matrices fixed by (ρ,σ) which have some pair of rows which are equal or some pair of columns which are equal.

4.4.4 Bringing it all together

Using Theorem 4.4.15 and Theorem 4.4.22 we can now determine

$$\left|\bigcup_{a \in \mathscr{A}(\rho,\sigma)} a\right| = \sum_{[A] \in P(\rho,\sigma)} k_{[A]} \left|\bigcap A\right|.$$

Recall that this union is the complement of the set $X_{m,n}^{\rho,\sigma}$ of all matrices fixed by (ρ,σ) with all rows and columns distinct in the set of all matrices fixed by (ρ,σ). We can use Lemma 4.3.1 to find the size of the set of matrices fixed by (ρ,σ). This will be two to the power of the number of orbits of $\langle(\rho,\sigma)\rangle$ in its action on $\mathbf{m} \times \mathbf{n}$, each orbit corresponds to a collection of entries which must have the same value and that value has two possibilities: 0 or 1. Let $c_1, c_2, \ldots$ be functions such that $c_i(g)$ returns the number of i cycles of the permutation g. Then the aforementioned number of orbits we seek is

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \gcd(i,j) * c_i(\rho) * c_j(\sigma)$$

and thus the number of elements of $X_{m,n}^{\rho,\sigma}$ is

$$\prod_{i=1}^{m}\prod_{j=1}^{n} 2^{\gcd\{i,j\}*c_i(\rho)*c_j(\sigma)}.$$

Therefore we can enumerate

$$|X_{m,n}^{(\rho,\sigma)}| = \prod_{i=1}^{m}\prod_{j=1}^{n} 2^{\gcd\{i,j\}*c_i(\rho)*c_j(\sigma)} - \left|\bigcup_{a\in\mathscr{A}(\rho,\sigma)} a\right|$$

as required. After determining all $|X_{m,n}^{(\rho,\sigma)}|$ for representatives (ρ,σ) of all conjugacy classes of $S_m \times S_n$ we insert these into the orbit counting formula as described in Section 4.2 and this gives us the number of $m \times n$ binary matrices with all rows unique and all columns unique up to row and column permutations.

4.4.5 Improvements

A formula for the number of $m \times n$ binary matrices with all rows unique and all columns unique is already known [39]. This is in fact a case we tackled during our enumeration, as this quantity is the same as the number of $m \times n$ binary matrices with all rows and all columns unique fixed by the pair of identity permutations $(1_{S_m}, 1_{S_n})$. However the known formula in this case is clearer and easier to apply so we would like to integrate it into our own method. We begin by presenting the formula together with a sketch of a proof. We will use the notation $s(n,k)$ to denote signed Stirling numbers of the first kind.

Lemma 4.4.23. *The number of $m \times n$ binary matrices with all rows unique and all columns unique is*

$$\sum_{i=1}^{m}\sum_{j=1}^{n} s(m,i)s(n,j)2^{ij}.$$

Proof. Let $E_{i,j}, F_{i,j}$ be the equivalences on $\mathbf{m} \times \mathbf{n}$ defined in Section 4.3, such that $\{0,1\}^{E_{i,j}}$ is the set of $m \times n$ binary matrices with row i equal to row j and $\{0,1\}^{F_{i,j}}$ contains the matrices with column i equal to column j. Then we apply this inclusion-exclusion principle on

$$I = \{\{0,1\}^{E_{i,j}} : 1 \leq i < j \leq m\} \cup \{\{0,1\}^{F_{i,j}} : 1 \leq i < j \leq n\}.$$

in order to enumerate the $m \times n$ binary matrices with all rows unique and all columns unique. The inclusion-exclusion principle yields

$$\left| \bigcup_{a \in I} a \right| = \sum_{A \subseteq I} (-1)^{|A|} |\cap A|. \tag{4.8}$$

Some consideration shows that subsets of I are in $1-1$ correspondence with graph pairs (G,H) where G has m vertices and H has n vertices. More precisely, if A is a subset of I then the corresponding graph pair has the edge (i,j) of G if and only if $E_{i,j} \in A$ and the edge (i,j) of H if and only if $F_{i,j} \in A$. Consider $I \subset I(\rho,\sigma)$ which corresponds with the graph pair (G,H). Then if vertex i,j are in the same connected component of G then row i and row j of matrices in $\cap A$ must be equal, and an analogous statement is true in terms of columns and H. If $X \subset \mathbf{m}$ and $Y \subset \mathbf{n}$ are the vertex sets of connected components of G and H, respectively, then all entries with index in $X \times Y$ of a matrix in I are equal. Therefore if there are $k(G)$ many connected components of G and $k(H)$ many connected components of H then the size of $|\cap A|$ is $2^{k(G)k(H)}$. Thus we may express the expression from Equation 4.8 as

$$\left| \bigcup_{a \in I} a \right| = \sum_{G} \sum_{H} (-1)^{E(G)+E(H)} 2^{k(G)k(H)} \tag{4.9}$$

where G sums over all graphs on $\mathbf{m}$ vertices and H sums over all graphs on $\mathbf{n}$ vertices; the function E returns the number of edges of a graph; and the function k returns the number of connected components of a graph. If we collect all terms with a the same number of connected components then we obtain

$$\left| \bigcup_{a \in I} a \right| = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(\sum_{\substack{G \text{ has } i \text{ connected} \\ \text{components}}} (-1)^{E(G)} \right) \left(\sum_{\substack{H \text{ has } j \text{ connected} \\ \text{components}}} (-1)^{E(H)} \right) 2^{ij}.$$

Then it has been shown[23] that the two innermost sums are equal to $s(m,i)$ and $s(n,j)$, where s returns the signed Stirling numbers of the first kind. This concludes the proof. □

We will again apply the inclusion-exclusion principle on the set $\mathscr{A}(\rho,\sigma)$ but we will ultimately use a different partition than the partition $P(\rho,\sigma)$ which we used before. Our method essentially considers a matrix fixed by (ρ,σ) in four parts. For any permutation $\rho \in S_m$ let $\mathrm{fix}(\rho)$ denote the subset of $\mathbf{m}$ fixed by ρ. Then define $\rho_=$ to be the identity permutation on $\mathrm{fix}(\rho)$ and ρ_* to be the permutation on $\mathbf{m} \setminus \mathrm{fix}(\rho)$ which acts identically to ρ on this set. Now we can consider a matrix fixed by some $(\rho,\sigma) \in S_m \times S_n$ as being composed of four sub-

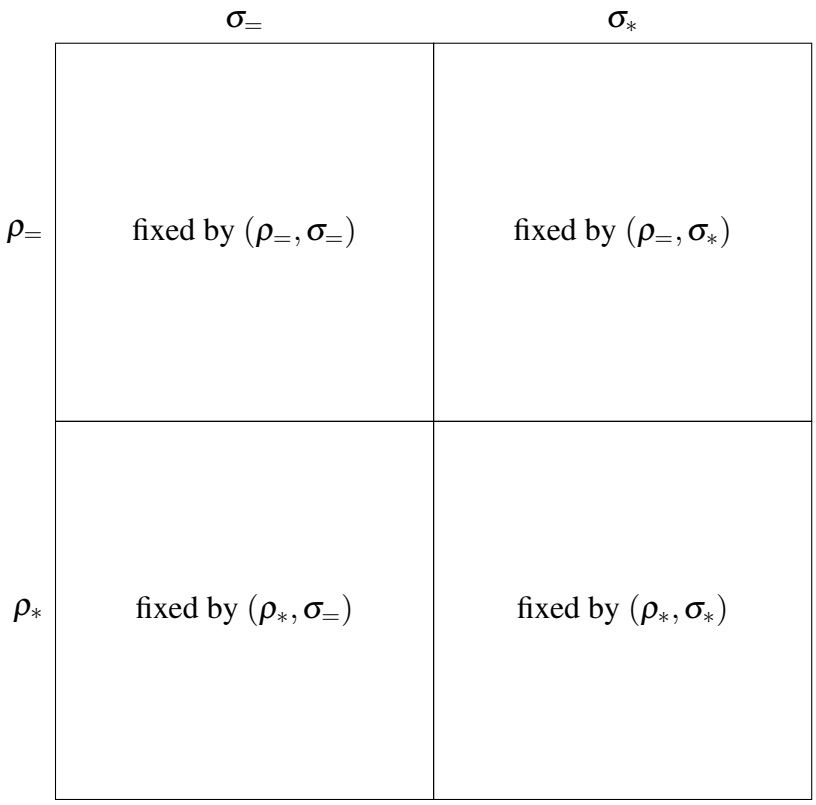


Fig. 4.8 A matrix fixed by (ρ, σ) can be seen as the composition of these four sub-matrices.

matrices, specifically those fixed by $(\rho_=, \sigma_=), (\rho_=, \sigma_*), (\rho_*, \sigma_=)$, and (ρ_*, σ_*). See Figure 4.8 for an illustration of this idea. Recall that

$$\left| \bigcup_{a \in \mathscr{A}(\rho,\sigma)} a \right| = \sum_{A \subseteq \mathscr{A}(\rho,\sigma)} (-1)^{|A|} \left| \bigcap A \right|. \tag{4.10}$$

Our approach begins by determining $|\cap A|$ in terms of the number of possibilities for the four sub-cases. Define

$$\mu_=(A) = \{i \in \mu(\mathbf{A}) : |\rho_{i,1}| = 1\},$$
$$\nu_=(A) = \{i \in \nu(\mathbf{A}) : |\sigma_{i,1}| = 1\},$$
$$\mu_*(A) = \mu(\mathbf{A}) \backslash \mu_=(A), \text{ and}$$
$$\nu_*(A) = \nu(\mathbf{A}) \backslash \nu_=(A).$$

So that $\mu_=(A)$ and $\nu_=(A)$ give the indices of the components of the graph pair corresponding to A which correspond with 1-cycles of ρ and σ, and $\mu_*(A), \nu_*(A)$ give the indices of the other components. Furthermore recall Ω from Equation 4.5 and define

$$\Omega_{=}(A) = 2^{|\mu_*(A)||\nu_1(A)|+|\mu_1(A)||\nu_*(A)|+|\mu_1(A)||\nu_1(A)|} \tag{4.11}$$

$$\Omega_*(A,\lambda) = \left(\prod_{(i,j)\in\mu_*(A)\times\nu_*(A)} \omega(\lambda(i,j))\right)\cdot\left(\prod_{i\in\mu_*(A)} \lambda(i,*)^{r_i(A)-1}\right)\cdot\left(\prod_{j\in\nu_*(A)} \lambda(*,j)^{s_j(A)-1}\right) \tag{4.12}$$

In a sense, $\Omega_*(A,\lambda)$ counts the number of possibilities for the sub-matrix fixed by (ρ_*,σ_*), whose sub-matrices corresponding to pairs of cycles have row periods described by λ. On the other hand, $\Omega_=(A)$ counts the combined number of possibilities for all three of the other sub-matrices, note that this does not depend on λ. If $\lambda\in\Lambda(A)$ then $\lambda(i,j)=1$ when $i\in\mu_=(A)$ or $j\in\nu_=(A)$ and so the values which λ takes on when evaluated in the expression for $\Omega_=$ are the same for all elements of $\Lambda(A)$. Next we show the relationship between $\Omega,\Omega_=$, and Ω_*.

Lemma 4.4.24. *Let $(\rho,\sigma)\in S_m\times S_n$. Let $A\subseteq\mathscr{A}(\rho,\sigma)$ and let $\lambda\in\Lambda(A)$. Then*

$$\Omega(A,\lambda) = \Omega_=(A)\cdot\Omega_*(A,\lambda).$$

Proof. Let $Y(A)$ denote the set $(\mu(\mathbf{A})\times\nu(\mathbf{A}))\setminus(\mu_*(A)\times\nu_*(A))$. By definition $\Omega(A,\lambda)/\Omega_*(A,\lambda)$ is equal to

$$\frac{\Omega(A,\lambda)}{\Omega_*(A,\lambda)} = \left(\prod_{(i,j)\in Y(A)} \omega(\lambda(i,j))\right)\cdot\left(\prod_{i\in\mu_=(A)} \lambda(i,*)^{r_i(A)-1}\right)\cdot\left(\prod_{j\in\nu_=(A)} \lambda(*,j)^{s_j(A)-1}\right)$$

For all $(i,j)\in Y(A)$ we have that $\lambda(i,j)=1$ since $\lambda(i,j)$ divides the greatest common divisor of $\delta_{R,i}(A)$ and $\delta_{C,j}(A)$ and either one or both of these must be 1 since $i\in\mu_=(A)$ or $j\in\nu_=(A)$. Therefore

$$\begin{aligned}\prod_{(i,j)\in Y(A)} \omega(\lambda(i,j)) &= \prod_{(i,j)\in Y(A)} \omega(1)\\ &= \prod_{(i,j)\in Y(A)} 2\\ &= 2^{|Y(A)|}\\ &= 2^{|\mu_*(A)||\nu_1(A)|+|\mu_1(A)||\nu_*(A)|+|\mu_1(A)||\nu_1(A)|}\\ &= \Omega_=(A).\end{aligned}$$

Finally we note that $\lambda(i,*)=1$ for all $i\in\mu_=(A)$ and $\lambda(*,j)=1$ for all $i\in\nu_=(A)$. The result follows. □

We now define a partition of $\mathscr{P}(\mathscr{A}(\rho,\sigma))$ such that we can apply the ideas of Lemma 4.4.23 to each part. Then instead of a sum across all $A \subset \mathscr{A}(\rho,\sigma)$ we may sum across all parts of the partition with a formula incorporating these ideas. Let $I(\rho,\sigma)$ denote the set

$$\{\{\{0,1\}^{E_{\rho_i,\rho_j}} : |\rho_i|,|\rho_j| = 1\} \cup \{\{0,1\}^{E_{\rho_i,\rho_j}} : |\sigma_i|,|\sigma_j| = 1\}.$$

In terms of the related graphs of subsets of $\mathscr{A}(\rho,\sigma)$ the set $I(\rho,\sigma)$ contains all the elements corresponding to edges between vertices that correspond to 1-cycles of ρ or σ. Note that these vertices are always labelled by the empty set. Therefore if we take some $A \subseteq \mathscr{A}(\rho,\sigma)$ and consider $A \backslash I(\rho,\sigma)$ then the corresponding graph pair has all vertices that correspond to 1-cycles having no edges and empty labels. If $A,B \subseteq \mathscr{A}(\rho,\sigma)$ are such that $A\backslash I(\rho,\sigma) = B\backslash I(\rho,\sigma)$ then we can say A and B are equal when we ignore the vertices corresponding to 1-cycles. We define the equivalence $\approx$ on $\mathscr{P}(\mathscr{A}(\rho,\sigma))$ by $A \approx B$ if and only if $A\backslash I(\rho,\sigma) = B\backslash I(\rho,\sigma)$. We now take a moment to demonstrate a property which is invariant amongst elements of the same class of the partition $\approx$. To do this we must define the function Λ_* for subsets of $\mathscr{A}(\rho,\sigma)$. Given A and $\lambda \in \Lambda(A)$ we define λ_* to be the restriction of λ to $\mu_*(A) \times \nu_*(A)$. Then $\Lambda_*(A)$ is defined to be the set

$$\Lambda_*(A) = \{\lambda_* : \lambda \in \Lambda(A)\}. \tag{4.13}$$

Essentially if we view $\lambda \in \Lambda(A)$ as a matrix then we discard the rows and columns corresponding to 1-cycles of ρ and σ to get λ_*. Now we can show a certain property is invariant amongst elements of a class of the partition $\approx$.

Lemma 4.4.25. *Let $(\rho,\sigma) \in S_m \times S_n$. Let $A \subseteq \mathscr{A}(\rho,\sigma)$. Then if $B \in [A]_\approx$ we have that*

$$\sum_{\lambda \in \Lambda(A)} \Omega_*(A,\lambda) = \sum_{\lambda \in \Lambda(B)} \Omega_*(B,\lambda)$$

Proof. By the definition of $\approx$ the graph pairs A and B are equal when removing the vertices corresponding to 1-cycles of ρ and σ. Let $\alpha : \mu_*(A) \to \mu_*(B)$ and $\beta : \nu_*(A) \to \nu_*(B)$ map the connected components (corresponding to non 1-cycles) of the graph pair corresponding to A to the identical ones in the graph pair corresponding to B. Then for all $\lambda \in \Lambda(A)$ there is a unique $\lambda' \in \Lambda(B)$ such that $\lambda(i,j) = \lambda'(i\alpha, j\beta)$ for all $(i,j) \in \mu_*(A) \times \nu_*(A)$. Moreover $\Omega_*(A,\lambda) = \Omega_*(B,\lambda')$ since the function $\Omega_*(A,\lambda)$ does not depend on the connected components of the graph pair corresponding to A which relate to 1-cycles of ρ or σ, and the same is true for B and λ'. Since the map $\lambda \mapsto \lambda'$ is a bijection from $\Lambda(A)$ to $\Lambda(B)$ the result follows immediately. □

Summing over $\approx$-classes

In this section we manipulate the expression given by applying the inclusion-exclusion principle on $\mathscr{A}(\rho,\sigma)$ in Equation 4.10. We will express the sum as a sum over the classes of the partition $\approx$ and recall Theorem 4.4.15 which gives us an expression for $|\cap A|$ using Ω.

$$\begin{aligned}\left|\bigcup_{a\in\mathscr{A}(\rho,\sigma)} a\right| &= \sum_{A\subseteq\mathscr{A}(\rho,\sigma)} (-1)^{|A|}\left|\bigcap A\right| \\ &= \sum_{[A]_\approx\subseteq\mathscr{A}(\rho,\sigma)}\ \sum_{B\in[A]_\approx} (-1)^{|B|}\left|\bigcap A\right| \\ &= \sum_{[A]_\approx\subseteq\mathscr{A}(\rho,\sigma)}\ \sum_{B\in[A]_\approx} (-1)^{|B|}\sum_{\lambda\in\Lambda(B)}\Omega(B,\lambda) \end{aligned} \tag{4.14}$$

Then we apply Lemma 4.4.24 and do some rearranging. Note that $\Omega_=(B)$ does not depend on λ so we can bring it outside the innermost sum. Then we apply Lemma 4.4.25 to bring the innermost sum outside the second sum since the value it takes is the same for all $B\in[A]_\approx$.

$$\begin{aligned}\left|\bigcup_{a\in\mathscr{A}(\rho,\sigma)} a\right| &= \sum_{[A]_\approx\subseteq\mathscr{A}(\rho,\sigma)}\ \sum_{B\in[A]_\approx} (-1)^{|B|}\sum_{\lambda\in\Lambda(B)}\Omega_=(B)\Omega_*(B,\lambda) \\ &= \sum_{[A]_\approx\subseteq\mathscr{A}(\rho,\sigma)}\ \sum_{B\in[A]_\approx} (-1)^{|B|}\Omega_=(B)\sum_{\lambda\in\Lambda(B)}\Omega_*(B,\lambda) \\ &= \sum_{[A]_\approx\subseteq\mathscr{A}(\rho,\sigma)}\left(\sum_{\lambda\in\Lambda(A)}\Omega_*(A,\lambda)\right)\cdot\sum_{B\in[A]_\approx}(-1)^{|B|}\Omega_=(B) \end{aligned} \tag{4.15}$$

Now we turn out attention to the innermost sum. Let $C=A\backslash I(\rho,\sigma)$ and note that $C=B\backslash I(\rho,\sigma)$ for any $B\in[A]_\approx$. The sum over $B\in[A]_\approx$ can be expressed as a sum over $I(\rho,\sigma)$ since each element of $[A]_\approx$ corresponds to the union of C with an element of $I(\rho,\sigma)$.

$$\begin{aligned}\sum_{B\in[A]_\approx}(-1)^{|B|}\Omega_=(B) &= \sum_{Y\subseteq I(\rho,\sigma)}(-1)^{|C\cup Y|}\Omega_=(C\cup Y) \\ &= (-1)^{|C|}\sum_{Y\subseteq I(\rho,\sigma)}(-1)^{|Y|}\Omega_=(C\cup Y) \end{aligned} \tag{4.16}$$

The set $I(\rho,\sigma)$ is in 1-1 correspondence with the set of all pairs of graphs (G,H) where G has $|\operatorname{fix}(\rho)|$ vertices and H has $|\operatorname{fix}(\sigma)|$. Therefore we can make the expression from Equation 4.16 look similar to the expression in Equation 4.9 the only difference being that $\Omega_=(C\cup Y)$ replaces

$2^{k(G)k(H)}$. Let $\Omega_=(G,H)$ denote the value of the $\Omega_=(C\cup Y)$ where Y is the element of $I(\rho,\sigma)$ corresponding to (G,H). Then

$$\sum_{Y\subseteq I(\rho,\sigma)}(-1)^{|Y|}\Omega_=(C\cup Y)=\sum_G\sum_H(-1)^{E(G)+E(H)}\Omega_=(G,H) \tag{4.17}$$

where the sums are over all graphs G with $|\operatorname{fix}(\rho)|$ vertices and H with $|\operatorname{fix}(\sigma)|$ vertices.

Now if $\Omega_=(G,H)$ is a function of the number of connected components of G and H, as was the case in Lemma 4.4.23 where we had $2^{k(G)k(H)}$ instead, then we can apply the method used in Lemma 4.4.23. Therefore let $Y\subseteq I(\rho,\sigma)$ and let (G,H) be the graph pair corresponding to Y. We have that $A\approx C\cup Y$ which implies $|\mu_*(A)|=|\mu_*(C\cup Y)|$ and $|\nu_*(A)|=|\nu_*(C\cup Y)|$, so these values do not depend on Y. Furthermore $|\mu_=(C\cup Y)|=k(G_Y)$ and $|\nu_=(C\cup Y)|=k(H_Y)$ so we have

$$\begin{aligned}\Omega_=(G,H)&=2^{|\mu_*(C\cup Y)||\nu_1(C\cup Y)|+|\mu_1(C\cup Y)||\nu_*(C\cup Y)|+|\mu_1(C\cup Y)||\nu_1(C\cup Y)|}\\&=2^{|\mu_*(A)|k(H_Y)+k(G_Y)|\nu_*(A)|+k(G_Y)k(H_Y)}.\end{aligned} \tag{4.18}$$

Thus $\Omega(G,H)$ depends only on C and the number of connected components of G and H. Finally, following the method of Lemma 4.4.23, we can deduce

$$\begin{aligned}&\sum_G\sum_H(-1)^{E(G)+E(H)}\Omega_=(G,H)\\&=\sum_{i=1}^{|\operatorname{fix}(\rho)||\operatorname{fix}(\sigma)|}\sum_{j=1}^{|\operatorname{fix}(\rho)||\operatorname{fix}(\sigma)|}\left(\sum_{\substack{G\text{ has }i\text{ connected}\\\text{components}}}(-1)^{E(G)}\right)\left(\sum_{\substack{H\text{ has }j\text{ connected}\\\text{components}}}(-1)^{E(G)}\right)2^{i|\nu_*(A)|+|\mu_*(A)|j+ij}\\&=\sum_{i=1}^{|\operatorname{fix}(\rho)||\operatorname{fix}(\sigma)|}\sum_{j=1}^{|\operatorname{fix}(\rho)||\operatorname{fix}(\sigma)|}s(|\operatorname{fix}(\rho)|,i),s(|\operatorname{fix}(\sigma)|,j)2^{i|\nu_*(A)|+|\mu_*(A)|j+ij}.\end{aligned} \tag{4.19}$$

For brevity we will denote

$$\chi(A)=\sum_{i=1}^{|\operatorname{fix}(\rho)||\operatorname{fix}(\sigma)|}\sum_{j=1}^{|\operatorname{fix}(\rho)||\operatorname{fix}(\sigma)|}s(|\operatorname{fix}(\rho)|,i),s(|\operatorname{fix}(\sigma)|,j)2^{i|\nu_*(A)|+|\mu_*(A)|j+ij}. \tag{4.20}$$

Now we can bring together our results. Returning to Equation 4.15 and substituting in Equation 4.16 then Equation 4.19 we obtain

$$\left|\bigcup_{a\in\mathscr{A}(\rho,\sigma)}a\right|=\sum_{[A]_\approx\subseteq\mathscr{A}(\rho,\sigma)}\left(\sum_{\lambda\in\Lambda(A)}\Omega_*(A,\lambda)\right)(-1)^{|A\setminus I(\rho,\sigma)|}\chi(A) \tag{4.21}$$

We can now evaluate the expression from the inclusion-exclusion principle by summing over $\approx$ classes rather than all elements of $\mathscr{A}(\rho,\sigma)$. However it seems as thought we have foregone our method of summing over the classes of the partition $P(\rho,\sigma)$ which we worked so hard to develop. The final part of this section involves combining these two approaches by summing over another partition called $\mathbb{P}(\rho,\sigma)$.

Summing over $\mathbb{P}$-classes

In this section we will again manipulate the expression given by applying the inclusion-exclusion principle on $\mathscr{A}(\rho,\sigma)$. We will express the sum as a sum over the classes of a new partition called $\mathbb{P}(\rho,\sigma)$ and utilize both the work on summing over $\approx$-classes and the earlier work of this chapter which involved summing over the classes of $P(\rho,\sigma)$. We define the partition $\mathbb{P}(\rho,\sigma)$ of $\mathscr{P}(\mathscr{A}(\rho,\sigma))$ to be such that the class of $A \subseteq \mathscr{A}(\rho,\sigma)$ is given by

$$[A]_{\mathbb{P}(\rho,\sigma)} = \{B \in \mathscr{A}(\rho,\sigma) : \exists C \in [A]_{P(\rho,\sigma)} \text{ such that } B \approx C\}.$$

Then the partition $\mathbb{P}(\rho,\sigma)$ can be seen to be the partition corresponding to the join of the equivalence $\approx$ with the equivalence corresponding to $P(\rho,\sigma)$. The key to combining the methods is thinking of every class of $\mathbb{P}(\rho,\sigma)$ as a union of $\approx$ classes.

$$\begin{aligned}\left|\bigcup_{a\in\mathscr{A}(\rho,\sigma)} a\right| &= \sum_{[A]_\approx \subseteq \mathscr{A}(\rho,\sigma)} \left(\sum_{\lambda\in\Lambda(A)} \Omega_*(A,\lambda)\right)(-1)^{|A\setminus I(\rho,\sigma)|}\chi(A) \\ &= \sum_{[A]_{\mathbb{P}(\rho,\sigma)}\in\mathbb{P}(\rho,\sigma)} \sum_{[B]_\approx \subseteq [A]_{\mathbb{P}(\rho,\sigma)}} \left(\sum_{\lambda\in\Lambda(B)} \Omega_*(B,\lambda)\right)(-1)^{|B\setminus I(\rho,\sigma)|}\chi(B) \qquad (4.22)\end{aligned}$$

We now work to tidy up the expression in Equation 4.22. First we demonstrate that if $[B]_\approx \subseteq [A]_{\mathbb{P}(\rho,\sigma)}$ then $\chi(A) = \chi(B)$. By looking at the definition of χ we see that this is true if $|\mu_*(A)| = |\mu_*(B)|$ and $|\nu_*(A)| = |\nu_*(B)|$. The cardinality of the sets returned by μ_* and ν_* are invariants of $P(\rho,\sigma)$ classes and of $\approx$ classes. Thus they are also invariants of $\mathbb{P}(\rho,\sigma)$ classes since it is the partition corresponding to the join of the other two equivalences.

Second, we will show that if $[B]_\approx \subseteq [A]_{\mathbb{P}(\rho,\sigma)}$ then

$$\left(\sum_{\lambda\in\Lambda(A)} \Omega_*(A,\lambda)\right) = \left(\sum_{\lambda\in\Lambda(B)} \Omega_*(B,\lambda)\right).$$

We must show that if $A \approx B$ or A, B are in the same class of $P(\rho, \sigma)$ then

$$\sum_{\lambda \in \lambda(A)} \Omega_*(A, \lambda) = \sum_{\lambda \in \lambda(B)} \Omega_*(B, \lambda).$$

The latter follows immediately from Lemma 4.4.25. To prove the former, recall that $P(\rho, \sigma)$ was defined such that if A, B are in the same class then $|\cap A| = |\cap B|$. Furthermore we have that

$$\left|\bigcap A\right| = \sum_{\lambda \in \lambda(A)} \Omega(A, \lambda) = \Omega_=(A) \sum_{\lambda \in \lambda(A)} \Omega_*(A, \lambda).$$

Since the cardinality of the sets returned by $\mu_=, \nu_=, \mu_*, \nu_*$ are invariants of classes of $P(\rho, \sigma)$ we have that $\Omega_=(A) = \Omega_=(B)$ and the result follows.

We now use these two results to improve Equation 4.22 to the following equation:

$$\left|\bigcup_{a \in \mathscr{A}(\rho,\sigma)} a\right| = \sum_{[A]_{\mathbb{P}(\rho,\sigma)} \in \mathbb{P}(\rho,\sigma)} \chi(A) \sum_{\lambda \in \Lambda(A)} \Omega_*(A, \lambda) \sum_{[B]_\approx \subseteq [A]_{\mathbb{P}(\rho,\sigma)}} (-1)^{|B \setminus I(\rho,\sigma)|}. \tag{4.23}$$

The final step will be to express the sum of parities

$$\sum_{[B]_\approx \subseteq [A]_{\mathbb{P}(\rho,\sigma)}} (-1)^{|B \setminus I(\rho,\sigma)|}$$

in terms of slightly modified versions of the functions γ, Ψ, Θ which were defined in Section 4.4.3. Recall Ψ from Equation 4.6 and define

$$\Psi_*(A) = \left[\prod_{x \in \mu_*(A)} \psi(r_i(A))\right] \cdot \left[\prod_{x \in \nu_*(A)} \psi(s_i(A))\right]. \tag{4.24}$$

Recall Θ from Equation 4.7 and define

$$\Theta_*(A) = \left(\prod_{x \in \mu_*(A)} \theta(|\rho_{i,1}|, \delta_{R,i}(A))\right) \cdot \left(\prod_{x \in \nu_*(A)} \theta(|\sigma_{i,1}|, \delta_{C,i}(A))\right). \tag{4.25}$$

Finally, recall γ from Lemma 4.4.16 and Lemma 4.4.17:

$$\gamma(A, \rho) = \frac{|C_{S_m}(\rho)|}{|Stab_{C_{S_m}((\rho))}(G_R(\bar{A}))|} = \frac{\prod_{i=1}^m a_i!}{\prod_{i=1}^m \prod_{j=1}^{a_i} (j!)^{k_{i,j}} \prod_{z=1}^m k_{i,j,z}!}.$$

Then we define

$$\gamma_*(A,\rho) = \frac{|C_{S_m}(\rho)|}{|Stab_{C_{S_m}((\rho))}(G_R(\bar{A}))|} = \frac{\prod_{i=2}^{m} a_i!}{\prod_{i=2}^{m}\prod_{j=1}^{a_i}(j!)^{k_{i,j}}\prod_{z=1}^{m} k_{i,j,z}!}, \tag{4.26}$$

and

$$\gamma_*(A) = \gamma_*(A,\rho)\gamma_*(A,\sigma). \tag{4.27}$$

Armed with these functions we are ready to state and prove the final technical result of this section.

Lemma 4.4.26. *Let* $(\rho,\sigma) \in S_m \times S_n$. *Let* $A \in \mathscr{A}(\rho,\sigma)$. *Then*

$$\sum_{[B]_\approx \subseteq [A]_{\mathbb{P}(\rho,\sigma)}} (-1)^{|B\setminus I(\rho,\sigma)|} = \gamma_*(A)\Psi_*(A)\Theta_*(A).$$

Proof. For every class $[B]_\approx$ we can find an element $C = B\setminus I(\rho,\sigma)$, such that $C \cap I(\rho,\sigma) = \emptyset$. Moreover this element is unique. Therefore there is a 1-1 correspondence between the sets

$$\{[B]_\approx : B \in [A]_{\mathbb{P}(\rho,\sigma)}\}, \text{ and } \{C \in \mathbb{P} : C \cap I(\rho,\sigma) = \emptyset\}.$$

In fact the set

$$\{C \in \mathbb{P} : C \cap I(\rho,\sigma) = \emptyset\}$$

is the class $[A\setminus I(\rho,\sigma)]_{P(\rho,\sigma)}$ of $P(\rho,\sigma)$. Therefore we can apply the theory of Section 4.4.3 to find the sum by parity of its elements.

$$\begin{aligned}\sum_{[B]_\approx \subseteq [A]_{\mathbb{P}(\rho,\sigma)}} (-1)^{|B\setminus I(\rho,\sigma)|} &= \sum_{B \in [A\setminus I(\rho,\sigma)]_{P(\rho,\sigma)}} (-1)^{|B|} \\ &= k_{[A\setminus I(\rho,\sigma)]_{P(\rho,\sigma)}} \\ &= \gamma(A\setminus I(\rho,\sigma))\Psi(A\setminus I(\rho,\sigma))\Theta(A\setminus I(\rho,\sigma)).\end{aligned} \tag{4.28}$$

All connected components of $A\setminus I(\rho,\sigma)$ containing vertices corresponding to 1-cycles of ρ and σ are of size 1 with empty vertex labels. By examining the definitions of γ,Ψ,Θ it is routine to confirm that $\gamma_*(A) = \gamma(A\setminus I(\rho,\sigma))$, $\Psi_*(A) = \Psi(A\setminus I(\rho,\sigma))$, and $\Theta_*(A) = \Theta(A\setminus I(\rho,\sigma))$. The result follows immediately. □

We end this section with by stating our improvements as a theorem. To do this we apply Lemma 4.4.26 to Equation 4.23.

Theorem 4.4.27. *Let* $(\rho,\sigma) \in S_m \times S_n$. *Then*

$$\left| \bigcup_{a \in \mathscr{A}(\rho,\sigma)} a \right| = \sum_{[A]_{\mathbb{P}(\rho,\sigma)} \in \mathbb{P}(\rho,\sigma)} \gamma_*(A)\Psi_*(A)\Theta_*(A)\chi(A) \sum_{\lambda \in \Lambda(A)} \Omega_*(A,\lambda).$$

Proof. This follows immediately from applying Lemma 4.4.26 to Equation 4.23. □

4.5 Results

The theory of this chapter has been implemented in GAP to enumerate $m \times n$ binary matrices with all rows unique and all columns unique up to row and column permutations. Herein we will describe these quantities as the number of *solutions* to the $m \times n$ case. Recall that a solution is a class of matrices which are equivalent up to row and column permutations. In this section we present the results of this enumeration along with some brief analysis. We begin by noting some relatively obvious facts about the results. First, the $m \times n$ and $n \times m$ cases are equivalent, so the result tables are symmetric about the main diagonal. Second, if $n > 2^m$ then there are no solutions in the $m \times n$ case (since there are 2^m distinct columns of length m). Next, if $n = 2^m$ then there is only 1 solution to the $m \times n$ case (a matrix in this class must have precisely all of the 2^m possible columns of length m). Finally, the $m \times 2^m - 1$ case has m solutions (each class corresponds to a choice of $2^m - 1$ of the 2^m possible columns). The result tables are shown below.

	n=1	n=2	n=3	n=4	n=5	n=6	n=7	n=8
m = 1	2	1	0	0	0	0	0	0
m = 2	1	3	3	1	0	0	0	0
m = 3	0	3	12	19	16	9	4	1
m = 4	0	1	19	94	250	459	649	729
m = 5	0	0	16	250	1796	8623	32016	98097
m = 6	0	0	9	459	8623	100494	881664	6363357
m = 7	0	0	4	649	32016	881664	17422636	277445249
m = 8	0	0	1	729	98097	6363357	277445249	9458731251
m = 9	0	0	0	655	255876	39482200	3772020128	271587692749
m = 10	0	0	0	477	577626	215647573	45164114945	6831373503141

	n=9	n=10	n=11	n=12
m = 4	655	477	284	136
m = 5	255876	577626	1139190	1974035
m = 6	39482200	215647573	1051483515	4619194764
m = 7	3772020128	45164114945	484983865296	4725468995643
m = 8	271587692749	6831373503141	153966229744772	3153375918715386
m = 9	16188344671676	838696116482956	38810836742429386	
m = 10	838696116482956	88664792374598854		

	n=13	n=14	n=15	n=16
m = 4	52	17	5	1
m = 5	3016854	4077039	4881084	5182325
m = 6	18405076501	66866590765	222481224532	680509104268
m = 7	42123674009470	345712057291064	2625688990951262	

	n=17	n=18	n=19	n=20	n=21	n=22	n=23
m = 5	4881092	4077077	3016994	1974438	1140090	579208	258092
m = 6	1919837989332						

	n=24	n=25	n=26	n=27	n=28	n=29	n=30	n=31	n=32
m = 5	100577	34230	10195	2674				33	1

If we plot the number of solutions for the $m \times n$ cases where $m = 3, 4$, or 5 and $1 \leq n \leq 2^m$ then the points appear to trace a bell curve which is slightly skewed, that is to say the gradient is slightly steeper on the left side of the peak than on the right side. If we plot the same data again but with a log scale on the y-axis then we see a parabola. The maximum number of solutions amongst $m \times n$ cases appears to always occur when $n = 2^{m-1}$.

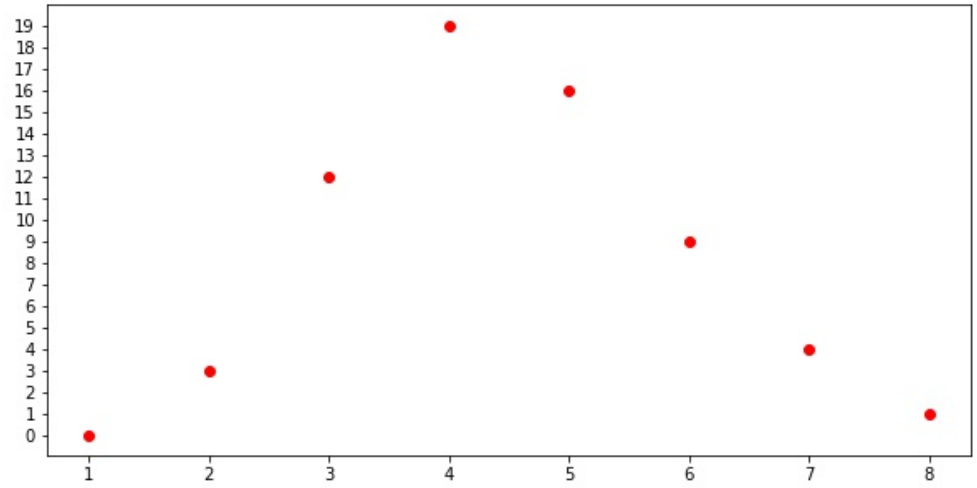


Fig. 4.9 The number of solutions to the $3 \times n$ cases.

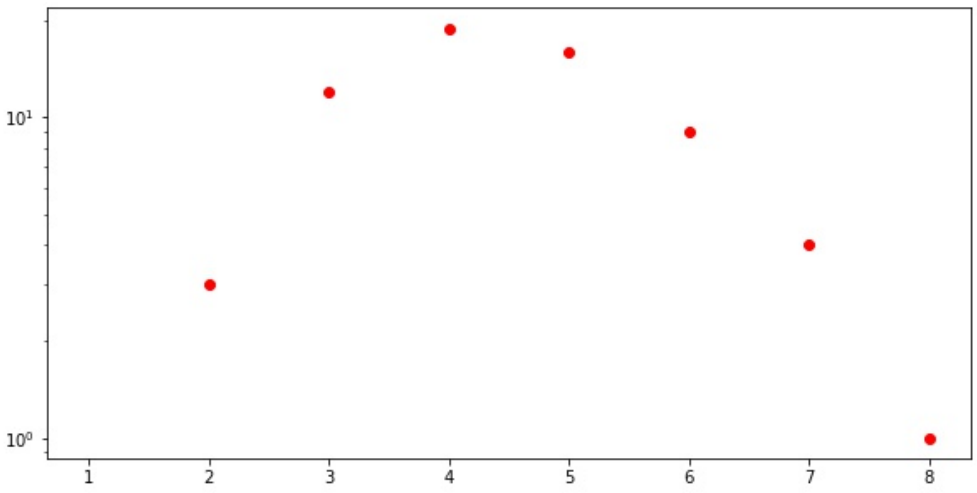


Fig. 4.10 The number of solutions to the $3 \times n$ cases.

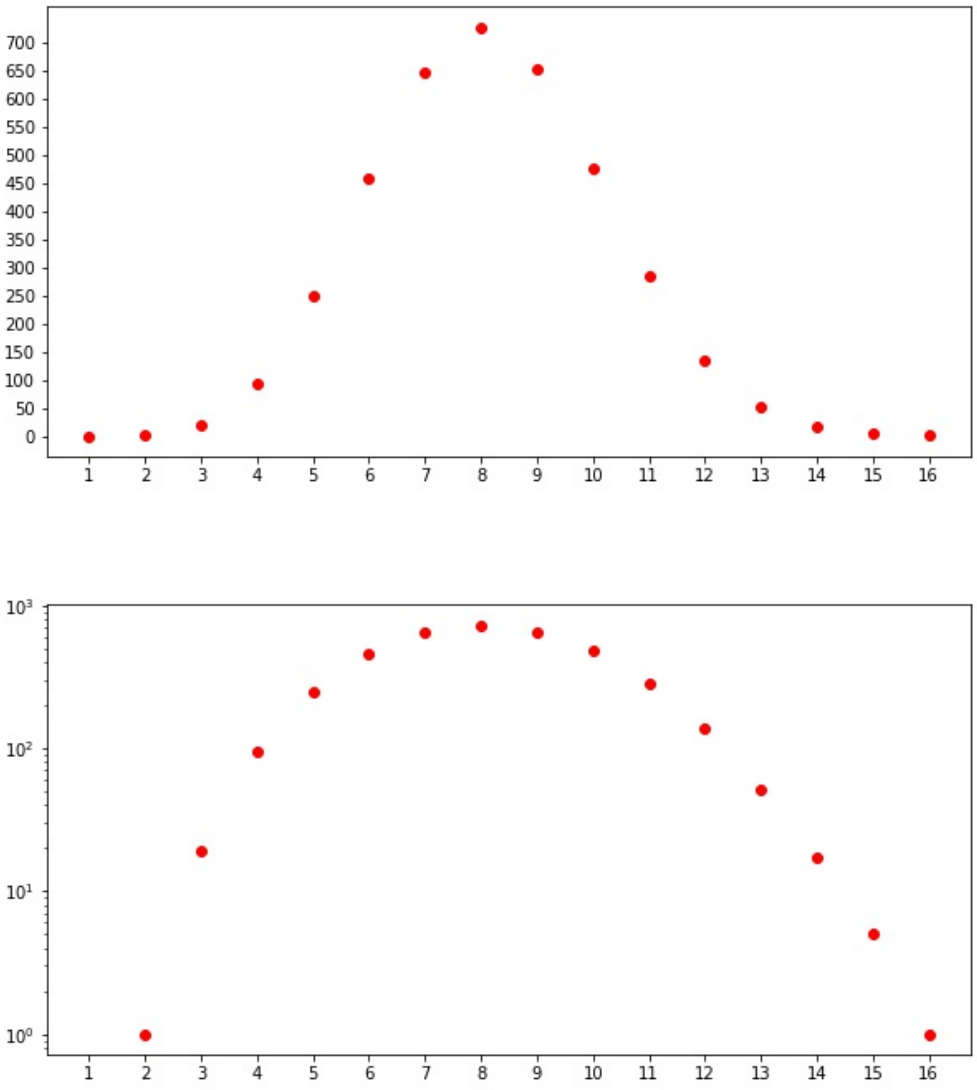


Fig. 4.11 The number of solutions to the $4 \times n$ cases.

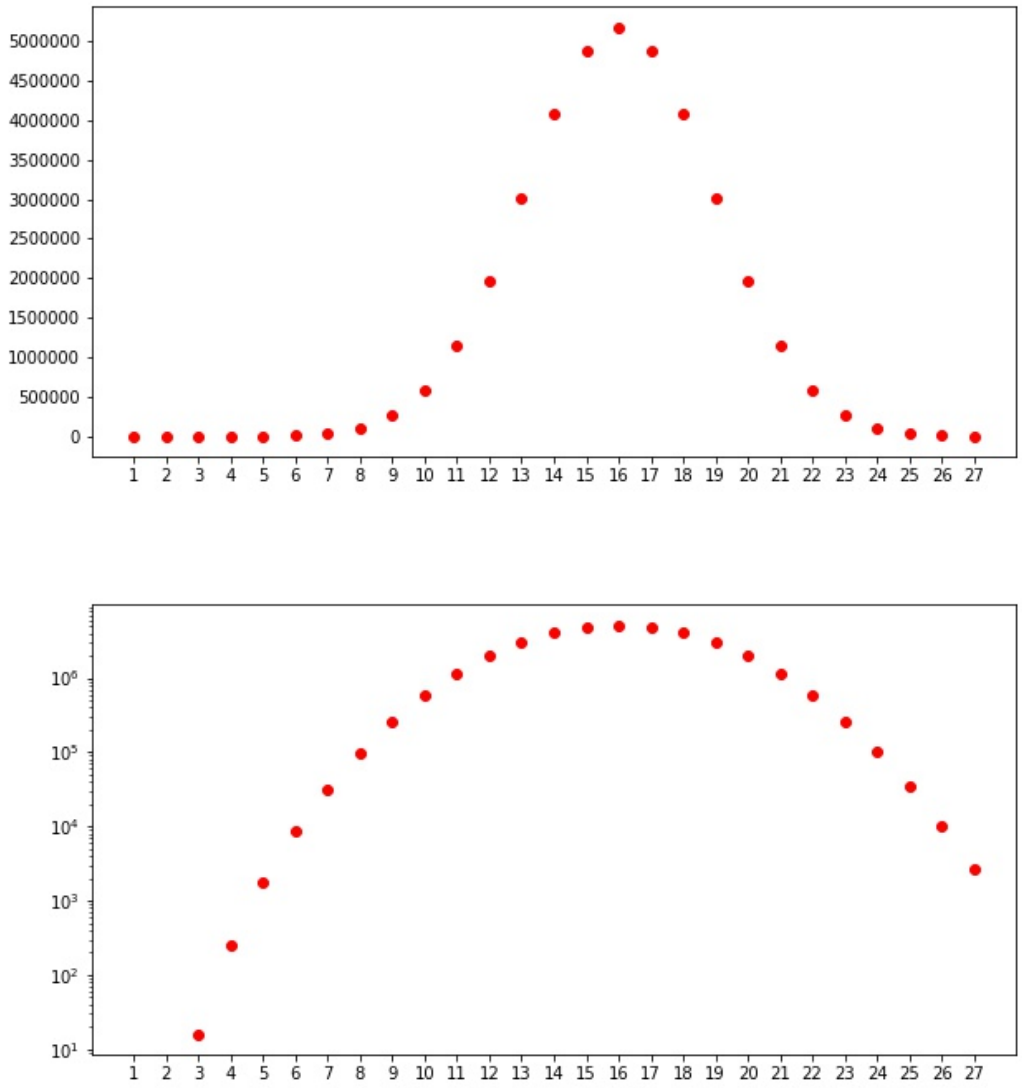


Fig. 4.12 The number of solutions to the $5 \times n$ cases, for $n \leq 27$.

4.5.1 Regular matrices

The motivation for enumerating these matrices stemmed from a desire to count isomorphism classes of congruence free semigroups. However, of the classes of matrices we counted, only those which are regular (no zero rows or zero columns) correspond to congruence free semigroups. Fortunately it is possible to determine the number of regular $m \times n$ binary matrices with all rows unique and all columns unique up to row and column permutation from the number of those which may or may not be regular. In this short section we will show how to do this and include tables showing the corresponding results.

Notice that a matrix with all rows distinct and all columns distinct can have at most one row of zeros and one column of zeros. A $m \times n$ binary matrix with all rows unique and all columns unique is of at least one of the following types:

1. regular,
2. has a row of zeros,
3. has a column of zeros, or
4. has both a row of zeros and a column of zeros.

Let $N(m,n)$ denote the set of equivalence classes of $m \times n$ binary matrices with all rows unique and all columns unique up to the equivalence of applying row and column permutations. To solve the $m \times n$ case we will need to have solved the $a \times b$ case for all pairs $(a,b) \in \mathbf{m} \times \mathbf{n} \backslash \{(m,n)\}$ as we will require information from these cases. This is not unreasonable since these cases should be easier to solve, for instance, we would not expect to be able to compute the 8×8 case if we can not compute the 7×7 case. Let $N_1(m,n)$, $N_2(m,n)$, $N_3(m,n)$, and $N_4(m,n)$ denote the subsets of $N(m,n)$ where the classes contain matrices of types 1, 2, 3, and 4, respectively. Then we have that $N_1(m,n)$ is the complement of $N_2(m,n) \cup N_3(m,n) \cup N_4(m,n)$ since a matrix is regular if and only if it contains no zero rows or columns. It follows that

$$|N_1(m,n)| = |N(m,n)| - |N_2(m,n) \cap N_3(m,n) \cap N_4(m,n)|. \tag{4.29}$$

Yet all matrices of type 4 are also matrices of type 2 and type 3 so $N_2(m,n) \cap N_3(m,n) = N_4(m,n)$. Therefore $|N_2(m,n) \cup N_3(m,n) \cup N_4(m,n)| = |N_2(m,n)| + |N_3(m,n)| - |N_4(m,n)|$. We also note that $N_2(m,n) = N_3(n,m)$. Furthermore, $|N_1(m-1,n-1)| = |N_4(m,n)|$ since if we remove the zero row and zero column from two matrices in the same class of $N_4(m,n)$ then we obtain two matrices in the same class of $N_1(m-1,n-1)$, and if we add a zero row and zero

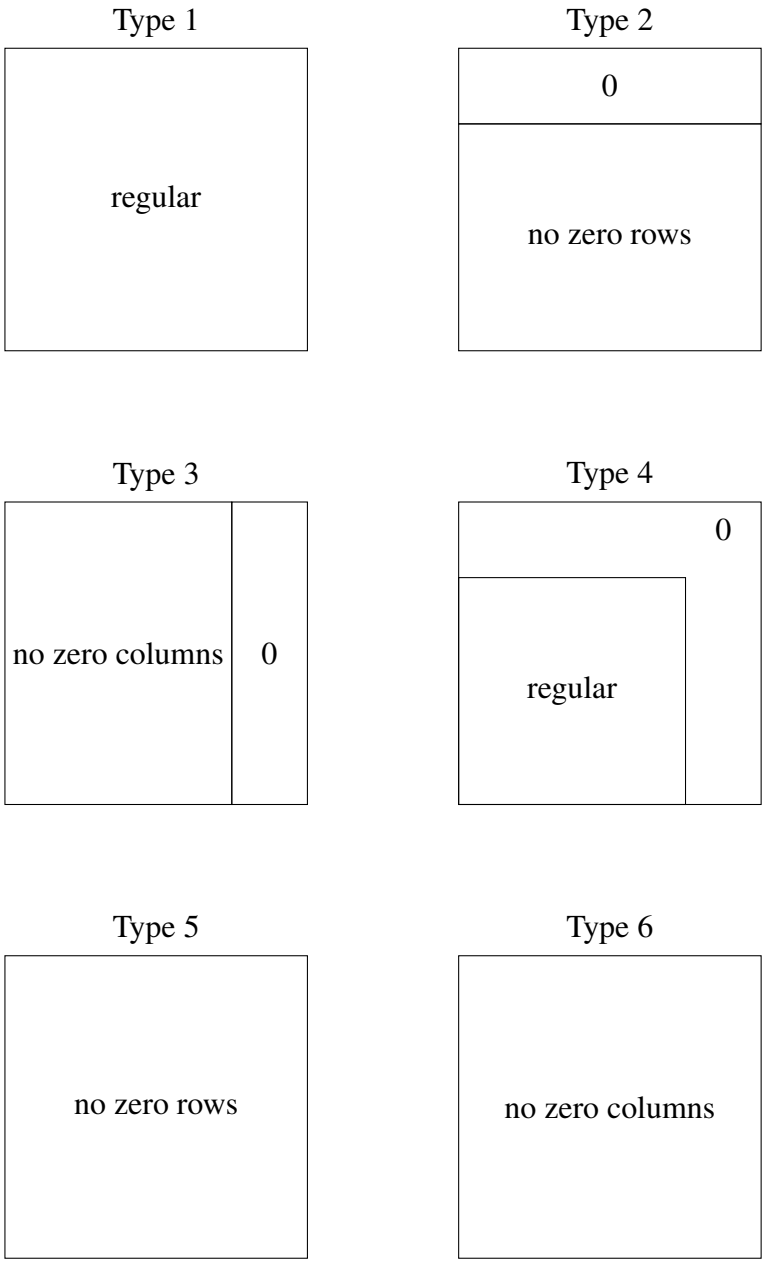


Fig. 4.13 Matrices of type 1, 2, 3, 4, 5, and 6.

column to two matrices in the same class of $N_1(m-1,n-1)$ then we obtain two matrices in the same class of $N_4(m,n)$. Therefore we may determine $|N_1(m,n)|$ by

$$\begin{aligned}|N_1(m,n)| &= |N(m,n)| - |N_2(m,n) \cap N_3(m,n) \cap N_4(m,n)| \\ &= |N(m,n)| - |N_2(m,n)| - |N_3(m,n)| + |N_4(m,n)| \\ &= |N(m,n)| - |N_2(m,n)| - |N_2(n,m)| + |N_1(m-1,n-1)| \end{aligned} \tag{4.30}$$

We assume that we have already solved the smaller dimension cases and therefore already know $|N_1(m-1,n-1)|$. It remains to determine $|N_2(m,n)|$ and $|N_(n,m)|$ in terms of the smaller dimension cases and $|N(m,n)|$, we also know. We will define two further types of binary matrix with all rows unique and all columns unique:

5. has no zero row,

6. has no zero column.

Matrices of types 1-6 are illustrated in Figure 4.13, although we will no longer worry about type 6 since they can be counted by the same method as for types 5, in the same way that types 2 and 3 are linked by transposition. A matrix of type 2 is formed of a row of zeros together with a $m-1 \times n$ matrix with all rows unique and all columns unique up to row and column permutations which has no zero rows, i.e. of type 5. It is easy to see that the classes of $m-1 \times n$ matrices of type 5 in $N(m-1,n)$ are in $1-1$ correspondence with the $m \times n$ matrices of type 2, up to permutations of rows and columns. If we let $N_5(m,n)$ denote the subset of $N(m,n)$ containing the classes of matrices of type 5 then

$$|N_2(m,n)| = |N_5(m-1,n)| \tag{4.31}$$

due to the correspondence previously mentioned. Furthermore we have that $N_2(m,n)$ and $N_5(m,n)$ are complements in $N(m,n)$ since a matrix either has a zero row or it does not. Thus we have

$$|N_5(m,n)| = |N(m,n)| - |N_2(m,n)|. \tag{4.32}$$

Repeated applications of Equations 4.31 and 4.32 then yields

$$\begin{aligned}
|N_2(m,n)| &= |N_5(m-1,n)| \\
&= |N(m-1,n)| - |N_2(m-1,n)| \\
&= |N(m-1,n)| - |N_5(m-2,n)| \qquad (4.33)\\
&= |N(m-1,n)| - |N(m-2,n)| + |N_2(m-2,n)| \\
&= |N(m-1,n)| - |N(m-2,n)| + |N_5(m-3,n)| \\
&= |N(m-1,n)| - |N(m-2,n)| + |N(m-3,n)| - |N_2(m-3,n)| \\
&\vdots \\
&= \sum_{i=1}^{m-1} (-1)^{i+1}|N(m-i,n)| + (-1)^{i+1}|N_2(1,n)|. \qquad (4.34)
\end{aligned}$$

Note that $|N_2(1,n)| = 1$ if and only if $n = 1$ and $|N_2(1,n)| = 0$ otherwise. Substituting Equation 4.34 into Equation 4.30 we obtain a formula for $|N_1(m,n)|$ in terms of $|N(a,b)|$ for lower dimension cases, $|N_2(1,m)|$, and $|N_2(1,n)|$. However if we have already solved the lower dimension cases we should already know $|N_2(m-1,n)|$ and can immediately evaluate 4.33 rather than using 4.34. Using the later method we may reformulate 4.30 to show that $|N_1(m,n)|$ is equal to:

$$|N(m,n)| - |N(m-1,n)| + |N_2(m-1,n)| - |N(n-1,m)| + |N_2(n-1,m)| + |N_1(m-1,n-1)|. \qquad (4.35)$$

We end this section by presenting tables of values for $|N_1(m,n)|$, which correspond to the number of isomorphism classes of congruence free semigroups which are Rees 0-matrix semigroups created from $m \times n$ binary matrices.

Chapter 4 Symbols

$X_{m,n}$ The subset of $\{0,1\}^{\mathbf{m}\times\mathbf{n}}$ containing only matrices which have all rows distinct and all columns distinct 84

$X_{m,n}^{(\rho,\sigma)}$ The subset of $X_{m,n}$ fixed by the action of $(\rho,\sigma) \in S_m \times S_n$. 85

$R_{\rho,\sigma}$ An equivalence relation defined on $\mathrm{dom}(\rho) \times \mathrm{dom}(\sigma)$ whose classes correspond to the orbits of the group generated by (ρ,σ). 85

dom Returns the domain of a function or permutation. 87

$\{0,1\}^P$ The subset of functions in $\{0,1\}^{\mathbf{m}\times\mathbf{n}}$ whose kernel contains the equivalence relation P. 89

$E_{x,y}$ An equivalence relation on $\mathbf{m}\times\mathbf{n}$ containing $((x,z),(y,z))$ for all $z \in \mathbf{n}$. 90

$F_{x,y}$ An equivalence relation on $\mathbf{m}\times\mathbf{n}$ containing $((z,x),(z,y))$ for all $z \in \mathbf{m}$. 90

$\rho_1,\ldots,\rho_r$ When ρ is a permutation these refer to the disjoint cycles of ρ, including 1-cycles. 92

$\sigma_1,\ldots,\sigma_s$ When σ is a permutation these refer to the disjoint cycles of σ, including 1-cycles. 92

r Often used to denote the number of disjoint cycles of a permutation called ρ, including 1-cycles. 92

s Often used to denote the number of disjoint cycles of a permutation called σ, including 1-cycles. 92

$E_{\rho_i,k}$ An equivalence relation on $\mathbf{m}\times\mathbf{n}$ equivalent to $E_{x,x\rho^k}$ for $x \in \rho_i$ where ρ_i is a disjoint cycle of ρ. 93

$F_{\sigma_j,k}$ An equivalence relation on $\mathbf{m}\times\mathbf{n}$ equivalent to $F_{x,x\sigma^k}$ for $x \in \mathrm{dom}(\sigma_i)$ where σ_i is a disjoint cycle of σ. 93

$\{0,1\}^{E_{\rho_i,\rho_j}}$ The union of all $\{0,1\}^{E_{x,y} \wedge R_{\rho,\sigma}}$ for $x \in \mathrm{dom}(\rho_i)$ and $y \in \mathrm{dom}(\rho_j)$. 96

$\{0,1\}^{F_{\sigma_i,\sigma_j}}$ The union of all $\{0,1\}^{F_{x,y} \wedge R_{\rho,\sigma}}$ for $x \in \mathrm{dom}(\sigma_i)$ and $y \in \mathrm{dom}(\sigma_j)$. 96

$f|_{I \times J}$ The sub-matrix of f on the domain $I \times J$. 98

$\mathscr{A}(\rho,\sigma)$ A collection matrix sets of the form $E_{\rho_i,k}, F_{\sigma_j,k}, E_{\rho_i,\rho_j}$, or F_{σ_i,σ_j}. 103

$\mathscr{P}$ The power set of a set X is denoted by $\mathscr{P}(X)$. 104

k_p A coefficient of Equation 4.4 which allows us to collect terms when applying the inclusion-exclusion principle. When p is a subset of $\mathscr{A}(\rho,\sigma)$, k_p equals the sum of the parities of the elements of p. 104, 125

$G(A)$ The graph pair $(G_R(A), G_C(A))$ associated with a subset A of $\mathscr{A}(\rho,\sigma)$. 104

$G_R(A)$ The row graph of the graph pair $G(A)$ associated with a subset A of $\mathscr{A}(\rho,\sigma)$. 104

$G_C(A)$ The column graph of the graph pair $G(A)$ associated with a subset A of $\mathscr{A}(\rho,\sigma)$. 104

$u_1,\ldots,u_r$ The vertices of a graph called $G_R(A)$ associated with a subset A of $\mathscr{A}(\rho,\sigma)$. 105

$v_1,\ldots,v_s$ The vertices of a graph called $G_C(A)$ associated with a subset A of $\mathscr{A}(\rho,\sigma)$. 105

$\mathscr{G}(\rho,\sigma)$ A collection of graph pairs in correspondence with the set $\mathscr{A}(\rho,\sigma)$. 106

K_R A function which maps subsets of $\mathscr{A}(\rho,\sigma)$ to the set of connected components of $G_R(A)$. 107

$\mu(A)$ The number of connected components of $G_R(A)$. 107

$\{K_{R,1}(A),\ldots,K_{R,\mu(A)}(A)\}$ Denotes the set $K_R(A)$ of connected components of $G_R(A)$. 107

$r_i(A)$ The size of the connected component $K_{R,i}(A)$ of the graph $G_R(A)$. 107

$\{u_{i,1},\ldots,u_{i,r_i(A)}\}$ The vertices of the connected component $K_{R,i}(A)$ of the graph $G_R(A)$. 107

$\rho_{i,j}$ The cycle of ρ corresponding to the vertex $u_{i,j}$ in the component $K_{R,i}(A)$ of the graph $G_R(A)$. 107

K_C A function which maps subsets of $\mathscr{A}(\rho,\sigma)$ to the set of connected components of $G_C(A)$. 107

$\nu(A)$ The number of connected components of $G_C(A)$. 107

$\{K_{C,1}(A),\ldots,K_{C,\nu(A)}(A)\}$ Denotes the set $K_C(A)$ of connected components of $G_C(A)$. 107

$s_i(A)$ The size of the connected component $K_{C,i}(A)$ of the graph $G_C(A)$. 108

$\{v_{i,1},\ldots,v_{i,s_i(A)}\}$ The vertices of the connected component $K_{C,i}(A)$ of the graph $G_C(A)$. 108

$\sigma_{i,j}$ The cycle of σ corresponding to the vertex $v_{i,j}$ in the component $K_{C,i}(A)$ of the graph $G_C(A)$. 108

L_A A function which maps vertices of the graphs $G_R(A)$ and $G_C(A)$ to their labels. 108

$\delta_R(A)$ The tuple $(\delta_{R,1}(A),\ldots,\delta_{R,\mu(A)}(A))$, see definition of $\delta_{R,i}$. 108

$\delta_{R,i}(A)$ The greatest common divisor of the labels of vertices in the connected component $K_{R,i}(A)$ of the graph $G_R(A)$ associated with a subset A of $\mathscr{A}(\rho,\sigma)$, or $|\rho_{i,1}|$ when the labels are all equal to the empty set. 108

$\delta_C(A)$ The tuple $(\delta_{C,1}(A),\ldots,\delta_{C,\nu(A)}(A))$, see definition of $\delta_{C,i}$. 108

$\delta_{C,i}(A)$ The greatest common divisor of the labels of vertices in the connected component $K_{C,i}(A)$ of the graph $G_C(A)$ associated with a subset A of $\mathscr{A}(\rho,\sigma)$, or $|\sigma_{i,1}|$ when the labels are all equal to the empty set. 108

$\bar{A}$ A subset of $\mathscr{A}(\rho,\sigma)$ constructed from the subset A of $\mathscr{A}(\rho,\sigma)$, see Lemma 4.4.5. 109

$C_{S_m\times S_n}((\rho,\sigma))$ The centralizer of (ρ,σ) in $S_m\times S_n$. 112

$P(\rho,\sigma)$ A partition of $\mathscr{A}(\rho,\sigma)$ where elements in the same class have the same size of intersection. 112

λ_f A map from $\{1,\ldots,\mu(A)\}\times\{1,\ldots,\nu(A)\}$ to $\mathbb{N}$ such that $\lambda_f(i,j)$ equals the row period of the sub-matrix $f|_{\mathrm{dom}(\rho_{i,1})\times\mathrm{dom}(\sigma_{j,1})}$. 114

$\Lambda(A)$ The set of all λ_f where $f\in\cap A$, for a subset A of $\mathscr{A}(\rho,\sigma)$. 114

Q_A The set of inverse images of the map from $\cap A$ to $\Lambda(A)$ which sends f to λ_f. 114

$\Omega(A,\lambda)$ The size of the set $\{f\in\cap A:\lambda_f=\lambda\}$, defined combinatorially in Equation 4.5. 116, 124

$\zeta(\rho,\sigma,\rho_i,\sigma_j,q)$ The set of all sub-matrices on the domain $\mathrm{dom}(\rho_i)\times\mathrm{dom}(\sigma_j)$ of matrices in $\{0,1\}^{R_{\rho,\sigma}}$, where the sub-matrix has row period equal to q. 117

$\omega(q)$ The size of the set $\zeta(\rho,\sigma,\rho_i,\sigma_j,q)$, which depends only on q 118

$\lambda(i,*)$ The least common multiple of the set of $\lambda(i,k)$ for $k \in \mu(A)$, where λ is some function in $\Lambda(A)$. 118

$\lambda(*,i)$ The least common multiple of the set of $\lambda(k,i)$ for $k \in \nu(A)$, where λ is some function in $\Lambda(A)$. 118

$\alpha_{i,j}$ Typical variable names for a set of constants used to enumerate matrices in a set $\cap A$, described in Lemma 4.4.12. 120

$\beta_{i,j}$ Typical variable names for a set of constants used to enumerate matrices in a set $\cap A$, described in Lemma 4.4.12. 120

$\gamma(A)$ The size of the set of all $\bar{B}$ such that B is in the $P(\rho,\sigma)$ class of A. Determined in Lemma 4.4.17. 126

$\psi(n)$ The sum over all connected graphs on n vertices by the parity of their number of edges. 128

$\Psi(A)$ The sum by edge parity of the graph pairs corresponding to elements of $\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}$ when $G_R(A)$ and $G_C(A)$ have no vertex labels. See also Lemma 4.4.19. 130

$\mathrm{div}(x)$ The set of proper divisors of the positive integer x. 131

$\theta(G,x,k)$ The sum by label parity of the graphs with vertex and edge set equal to G, labels which are subsets of the proper divisors of x, and greatest common divisor of the union of the labels equal to k. We define $\theta(G,x,x) = 1$ to refer to the case where all labels are the empty set. See also Lemma 4.4.20. 131

$\theta(x,k)$ Equal to $\theta(G,x,k)$, which was demonstrated to not depend on G in Lemma 4.4.20. 132

$\mathscr{B}(A)$ The subset of $\{B \subseteq \mathscr{A}(\rho,\sigma) : \bar{B} = \bar{A}\}$ containing only those elements whose graph pair has the same edge sets as the graph pair of A. 133

$s(n,k)$ Signed Stirling number of the first kind. 137

$\rho_=$ The restriction of the permutation ρ to the domain $\mathrm{fix}(\rho)$. 138

ρ_* The restriction of the permutation $\rho \in S_m$ to the domain $\mathbf{m} \setminus \mathrm{fix}(\rho)$. 138

$\mu_=(A)$ The subset of $\{1,\ldots,\mu(A)\}$ such that $i \in \mu(A)$ if and only if $\rho_{i,1}$ is a 1-cycle. 139

$\nu_=(A)$ The subset of $\{1,\ldots,\nu(A)\}$ such that $i \in \nu(A)$ if and only if $\sigma_{i,1}$ is a 1-cycle. 139

$\mu_*(A)$ The set $\{1,\ldots,\mu(A)\} \setminus \mu_=(A)$. 139

$\nu_*(A)$ The set $\{1,\ldots,\nu(A)\}\backslash\nu_=(A)$. 139

$\Omega_=(A)$ See Equation 4.11. 140

$\Omega_*(A,\lambda)$ See Equation 4.12. 140

$I(\rho,\sigma)$ The subset of $\mathscr{A}(\rho,\sigma)$ containing only: E_{ρ_i,ρ_j} where ρ_i,ρ_j are 1-cycles, and F_{σ_x,σ_y} where σ_i,σ_j are 1-cycles. 141

$\approx$ The equivalence on $\mathscr{P}(\mathscr{A}(\rho,\sigma))$ such that $A\approx B$ if and only if $A\backslash I(\rho,\sigma)$ is equal to $B\backslash I(\rho,\sigma)$. 141

λ_* The restriction of $\lambda\in\Lambda(A)$ to the domain $\mu_*(A)\times\nu_*(A)$, where A is some subset of $\mathscr{A}(\rho,\sigma)$. 141

$\Lambda_*(A)$ The set containing all λ_* where λ is in $\Lambda(A)$. 141

$\chi(A)$ See Equation 4.20. 143

$\mathbb{P}(\rho,\sigma)$ The partition corresponding to the join of the equivalences $\approx$ and $P(\rho,\sigma)$. 144

$\Psi_*(A)$ See Equation 4.24. 145

$\Theta_*(A)$ See Equation 4.25. 145

$\gamma_*(A,\rho)$ See Equation 4.26. 146

$\gamma_*(A)$ See Equation 4.27. 146

Chapter 5

E-unitary inverse semigroups

An admiration for symmetry seems part of human nature and it is therefore not surprising that group theory is one of the most successful branches of algebra. However many authors have come to appreciate the role of partial symmetries in tackling problems of symmetry. Inverse semigroups are to partial symmetries what groups are to symmetries and consequentially are an area of interest to many.

Within the class of inverse semigroups the E-unitary inverse semigroups are of great importance [29, §7]. Perhaps the most famous result relating to E-unitary inverse semigroups is the McAlister covering theorem which shows that every inverse semigroups is a finite idempotent-separating homomorphic image of an E-unitary inverse semigroup, and it provides ample motivation for their study. Furthermore the study of E-unitary inverse semigroups has been greatly enhanced by the discovery of their representation by McAlister triple semigroups, in a similar vein to the benefit provided to the study of 0-simple semigroups by the discovery of the Rees 0-matrix semigroup representation. Finally, we note that many important examples of inverse semigroups are E-unitary such as the bicyclic monoid, semidirect products of semilattices by groups, groups, and free inverse semigroups.

The aim of this chapter is to compare E-unitary inverse covers. We are motivated by the question of finding the 'best' E-unitary inverse cover of an inverse semigroup up to isomorphism. This requires a notion of 'best' such as being minimal with respect to a sensible ordering.

The structure of this chapter is as follows. In Section 5.1 we introduce the necessary background theory of inverse semigroups, E-unitary inverse semigroups, and E-unitary inverse covers, as well as Clifford semigroups which will provide an important, yet relatively simple, class of examples. In Section 5.2 we define two orderings that seem natural for comparing E-unitary inverse covers. Essentially these two orderings are: (i) $P \leq_1 Q$ if there exists a surjective homomorphism from Q to P, and (ii) $P \leq_2 Q$ if there exist an injective homomorphism from P into Q. Section 5.3 examines E-unitary inverse covers of various Clifford semigroups and

prove results about their E-unitary inverse covers. Furthermore, examples provided in this section show that the two orderings defined in the previous section are unsatisfactory, that is to say, they are unable to provide a unique minimal cover in general. Finally, in Section 5.4 we pose conjectures about E-unitary inverse covers of Clifford semigroups based on the examples of the previous section. In particular, we suggest a weaker equivalence than isomorphism for comparing covers, and conjecture that, up to this equivalence, we may be able to find a minimal cover for certain Clifford semigroups.

5.1 Preliminaries

There are several equivalent ways to define inverse semigroups. The most common are summarised in the following theorem.

Theorem 5.1.1. *[19, Theorem 5.1.1] Let S be a semigroup. Then the following statements are equivalent:*

(i) *S is an inverse semigroup;*

(ii) *S is regular, and its idempotents commute;*

(iii) *every $\mathscr{L}$-class and every $\mathscr{R}$-class contains exactly one idempotent;*

(iv) *every element of S has a unique inverse.*

Let X and Y be sets. A *partial function* from X to Y is a function from a subset of X to a subset of Y. Of greatest interest to us are those partial functions which are bijections. The collection of partial bijections from X to X together with the operation of composition is known as the symmetric inverse monoid, $\mathscr{I}_X$. We will write $\mathscr{I}_n$ to denote the symmetric inverse monoid on the set $\{1,2,\ldots,n\}$. The semigroup $\mathscr{I}_X$ is to inverse semigroups what the full transformation monoid is to semigroups, or what the symmetric group is to groups.

Theorem 5.1.2 (Vagner-Preston Representation Theorem)**.** *[35, 40] For any inverse semigroup S there exists an embedding into the symmetric inverse monoid $I(S)$.*

Next we define E-unitary inverse semigroups, which are the focus of this chapter.

Definition 5.1.3. Let S be an inverse semigroup. We say that S is E-unitary if for all e in $E(S)$ and s in S:

$$es \in E \implies s \in E,$$

or equivalently [19, §5.9]:

$$se \in E \implies s \in E.$$

Before we can continue discussing E-unitary inverse semigroups we must discuss semigroup homomorphisms between inverse semigroups. First we recall how semigroup homomorphisms behave on inverse semigroups.

Theorem 5.1.4. *[19, Theorem 5.1.4] Let S be an inverse semigroup, let T be a semigroup and let ϕ be a semigroup homomorphism from S to T. Then* $\operatorname{im}\phi$ *is an inverse semigroup. Moreover, ϕ is an inverse semigroup morphism.*

There are two properties of congruences which are key to the study of inverse semigroups congruences, the trace and the kernel.

Definition 5.1.5. Let S be an inverse semigroup and let ρ be a congruence on S. The restriction of ρ to $E(S)$ is called the *trace* of ρ and is denoted by $\operatorname{tr}\rho$. The union of all ρ-classes containing idempotents is called the *kernel* of ρ and is denoted by $\operatorname{Ker}\rho$.

The trace of an inverse semigroup S is a congruence on $E(S)$ and the kernel of S is a subsemigroup. We now define a property of congruences on $E(S)$ which the trace satisfies and a property of subsemigroups of S which the kernel satisfies. The aim is to show a correspondence between these special congruences of $E(S)$ and the possible traces, as well as a correspondence between these special subsemigroups and the possible kernels.

Proposition 5.1.6. *[19, §5.3] Let S be an inverse semigroup and let ρ be a congruence on S. Then*

(i) $\tau = \operatorname{tr}\rho$ *is* normal *in the sense that*

$$e\tau f \implies (\forall s \in S)\ s^{-1}es\tau s^{-1}fs;$$

(ii) $N = \operatorname{Ker}\rho$ *is a* normal *subsemigroup of S in the sense that*

$$a \in N \implies (\forall s \in S)\ s^{-1}as \in N.$$

The trace and kernel of an inverse semigroup congruence are not independent. We now define congruence pairs, which are essentially those trace and kernel pairs which correspond to a congruence.

Definition 5.1.7. Let S be an inverse semigroup. If τ is a normal congruence on $E(S)$ and N is a normal subsemigroup of S then we will call the pair (N,τ) a *congruence pair* of S if for all s in S and e in E:

(i) $se \in N$ and $(e, s^{-1}s) \in \tau \implies s \in N$;

(ii) $s \in N \implies (ss^{-1}, s^{-1}s) \in \tau$.

The following theorem shows the correspondence between inverse semigroup congruences and congruence pairs.

Theorem 5.1.8. *[19, Theorem 5.3.3] Let S be an inverse semigroup. If ρ is a congruence on S then $(\operatorname{Ker}\rho, \operatorname{tr}\rho)$ is a congruence pair. Conversely, if (N,τ) is a congruence pair, then the relation*

$$\rho_{(N,\tau)} = \{(s,t) \in S \times S : (s^{-1}s, t^{-1}t) \in \tau, st^{-1} \in N\}$$

is a congruence on S. Moreover $\operatorname{Ker}\rho_{(N,\tau)} = N$, $\operatorname{tr}\rho_{(N,\tau)} = \tau$, and $\rho_{(\operatorname{Ker}\rho, \operatorname{tr}\rho)} = \rho$.

There is a special congruence called the minimum group congruence which has the property of being the least congruence on an inverse semigroup S such that the quotient is a group. Furthermore any other congruence $\rho \subset S \times S$ such that S/ρ is a group contains the minimum group congruence.

Definition 5.1.9. The *minimum group congruence* σ is defined on the inverse semigroup S by

$$s \sigma t \text{ if and only if } \exists u \leq s,t$$

for all $s,t \in S$.

The minimum group congruence will play a key role in refining the idea of an E-unitary cover of an inverse semigroup. An *E-unitary cover* of an inverse semigroup S is an E-unitary inverse semigroup P such that there exists a surjective idempotent separating homomorphism $\theta : P \to S$. The term *idempotent separating* indicates that if $e, f \in E(P)$ then $e\theta = f\theta$ implies $e = f$. Alternatively, θ is an idempotent separating homomorphism if $\operatorname{tr}\theta$ equals the diagonal relation on $E(P)$. The following result is the famous McAlister covering theorem [32]. We include an abridged proof as we will later require a construction involved in the proof. Note that a *factorisable inverse monoid* is an inverse monoid M such that $M = E(M)U(M)$, that is every element of M can be written as a product eg of an idempotent e with an element g of the group of units $U(M)$.

Theorem 5.1.10 (McAlister Covering Theorem)**.** *Every (finite) inverse semigroup is an idempotent separating homomorphic image of a (finite) E-unitary inverse semigroup.*

Proof sketch. Let S be a (finite) inverse semigroup then S may be embedded in a (finite) symmetric inverse monoid $I(X)$ by the Vagner-Preston representation theorem. If X is a finite set then the symmetric inverse monoid $I(X)$ is a finite factorisable inverse monoid. Otherwise, if X is an infinite set then $I(X)$ can be embedded in a factorisable inverse monoid, and thus any

semigroup which can be embedded in $I(X)$ can be embedded in a factorisable inverse monoid by composing the two embeddings.

Let $\iota : S \to F$ be an embedding of S into a factorisable inverse monoid F, and let $U(F)$ denote the group of units of F. Set

$$P = \{(s,g) \in S \times U(F) : \iota(s) \leq g\}.$$

It can be shown that P is an E-unitary inverse semigroup. Now if we define $\theta : P \to S$ by $(s,g)\theta = s$ then θ is a surjective idempotent separating homomorphism, as required. □

The construction of the cover in Theorem 5.1.10 relies on finding an embedding into a factorisable inverse monoid. Lawson has since provided a refinement of this result. To understand this refinement we first refine the notion of an embedding ι of an inverse semigroup S into a factorisable inverse monoid F. Such an embedding $\iota : S \to F$ is said to be *strict* if for all $g \in U(F)$ there exists $s \in S$ such that $(s)\iota \leq g$. Roughly speaking, the choice of F in a strict embedding is efficient in the sense that $U(F)$ is not unnecessarily large. When such an embedding $\iota : S \to F$ is not strict it is easy to construct an embedding which is strict. To do this, let:

$$H = \{g \in U(F) : (s)\iota \leq g \text{ for some } s \in S\}$$

then H is a subgroup of $U(F)$ and

$$[H] = \{f \in F : (\exists h \in H)(f \leq h)\}$$

is a factorisable inverse submonoid of F with group of units H. Then the embedding $\iota : S \to [H]$ is strict. We call $[H]$ the *closure of H in F*. Now we apply the construction of Theorem 5.1.10 using a strict embedding and obtain the following result.

Theorem 5.1.11. *[29, §7 Theorem 7] Let S be an inverse semigroup, and let $\iota : S \to F$ be a strict factorisable embedding. Let P be the E-unitary cover of S constructed from the embedding by the method of Theorem 5.1.10. Then P/σ is isomorphic to $U(F)$.*

Using this result Lawson refines the notion of an E-unitary cover. Let P be an E-unitary inverse semigroup and suppose that there is a surjective idempotent separating homomorphism $\theta : P \to S$. Then we say that P is an *E-unitary cover of S over P/σ*, where σ is the minimum group congruence of P. A strict factorisable embedding $\iota : S \to F$ will thus give rise to an E-unitary cover of S over a group isomorphic to $U(F)$. Lawson's refinement leads to the following result which shows the correspondence between E-unitary covers over groups and strict embeddings into factorisable inverse monoids.

Theorem 5.1.12. *[29, §8 Theorem 10] Every E-unitary cover of an inverse semigroup S over a group G is isomorphic to one constructed from a strict embedding of S into a factorisable inverse monoid with group of units isomorphic to G.*

We now introduce Clifford semigroups, which will be an important source of examples of E-unitary inverse covers of inverse semigroups.

Definition 5.1.13. An inverse semigroup S is a *Clifford semigroup* if and only if every $\mathscr{D}$-class is a group.

Clifford semigroups are conveniently represented by semilattices of groups. We believe this representation lends much intuition to the analysis of these semigroups.

Definition 5.1.14. Let E be a semilattice. Let $G = \{G_e : e \in E\}$ be a set of groups indexed by E. Let $\phi = \{\phi_{e,f} : G_e \to G_f \mid e, f \in E \text{ and } e \geq f\}$ be a set of group homomorphisms which are compatible, in the sense that if $d \geq e \geq f$ are elements of E then the composition $\phi_{d,e} \circ \phi_{e,f}$ is equal to $\phi_{d,f}$. Then the *semilattice of groups* $S = (E, G, \phi)$ is a semigroup with elements

$$\bigcup_{e \in E} G_e$$

and the product of $g \in G_e$ and $h \in G_f$ is

$$gh = (g)\phi_{e,ef}(h)\,\phi_{f,ef}$$

which is an element G_{ef}.

Finally, the last result of this section allows us to speak of Clifford semigroups and semilattices of groups interchangeably.

Proposition 5.1.15. *[19, Theorem 4.2.1] A semigroup is isomorphic to a semilattice of groups if and only if it is a Clifford semigroup.*

5.2 Comparing covers

Knowing that for every finite inverse semigroup there is a finite E-unitary semigroup which covers it, a natural question to ask is: which is the 'best' cover? Or more generally: when is one cover 'better' than another? As seen in Theorem 5.1.12, Lawson seems to take a step towards this goal with his refinement of the notion of an E-unitary cover to that of an E-unitary cover over a group. We will present several novel notions of how to compare covers in this section.

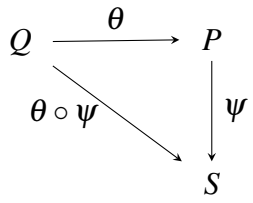


Fig. 5.1 The cover of S by P is superior than or equivalent to the cover of S by Q according to the composition ordering.

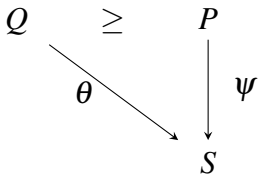


Fig. 5.2 The cover of S by P seems superior than or equivalent to the cover of S by Q.

Later we will show how each of these notions fall short and give examples demonstrating that, in general, a unique 'best' cover does not exist for any sensible notion of comparison.

Let S be an inverse semigroup. Furthermore, let P, Q be E-unitary inverse semigroups such that the maps $\psi : P \to S$ and $\theta : Q \to P$ are surjective idempotent separating homomorphisms. Then S is also covered by Q via $\theta \circ \psi$. It seems natural to the author to say that P is the superior cover of S (unless θ is an isomorphism, in which case both covers are isomorphic). This situation is displayed in Figure 5.1. We will denote this relationship by $P \leq_c Q$ and we call this relation the *composition order* on covers of S.

The author also considers the case where S is an inverse semigroup which is covered by E-unitary inverse semigroups P, Q such that P is a subsemigroup of Q (Figure 5.2). In this case it seems natural to say that Q is the superior cover. More generally, if instead of P being a subsemigroup of Q we have that P embeds into Q the author suggests that ψ is the superior cover (Figure 5.3). We will denote this relationship by $\theta \leq_e \psi$ and we call this relation the *embedding order* on covers of S.

Now that we have introduced our methods of comparing covers we will define them explicitly. For each inverse semigroup S we take the collection of all covers of S up to isomorphism and define relations on this set corresponding to the comparison order and the embedding order. We then will go on show that each of these relations possess the desirable quality of being a partial order.

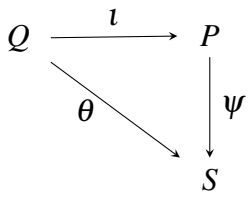


Fig. 5.3 Let ι be an embedding. Then the cover of S by P is superior than or equivalent to the cover of S by Q according to the embedding ordering.

Definition 5.2.1. Let S be an inverse semigroup and let X denote the collection of all E-unitary inverse covers of S, up to isomorphism. Let P,Q be E-unitary inverse covers of S, and let $[P],[Q] \in X$ denote the collections of all semigroups isomorphic to P,Q, respectively. Then define the relations $\leq_c, \leq_e$ on X as follows:

(i) $[P] \leq_c [Q]$ if and only if there exists a surjective idempotent separating homomorphism $\theta : Q \to P$, in other words Q is an E-unitary inverse cover of P;

(ii) $[P] \leq_e [Q]$ if and only if there exists a monomorphism $\theta : P \to Q$, in other words if P embeds into Q.

The composition ordering is now shown to be a partial order.

Proposition 5.2.2. *Let S be an inverse semigroup and let X denote the collection of all E-unitary inverse covers of S, up to isomorphism. Then $\leq_c$ is a partial order.*

Proof. We will write P,Q,R to denote E-unitary inverse covers of S, and will write $[P],[Q],[R] \in X$ denote the collections of all semigroups isomorphic to P,Q,R, respectively. For any semigroup the identity mapping is an idempotent separating surjective homomorphism. Therefore $[P] \leq_c [P]$ for any $[P] \in X$, and so $\leq_c$ is reflexive. Now assume that $[P],[Q] \in X$ such that $[P] \leq_c [Q]$ and $[Q] \leq_c [P]$. Then there must exist $\psi_1 : Q \to P$ and $\psi_2 : P \to Q$ which are surjective idempotent separating homomorphisms. Since ψ_1, ψ_2 are both surjective P and Q must be isomorphic. Therefore $[P]$ equals $[Q]$ and $\leq_c$ is anti-symmetric. Next assume that $[P],[Q],[R] \in X$ satisfy $[P] \leq_c [Q]$ and $[Q] \leq_c [R]$. Then there exists $\psi_1 : R \to Q$ and $\psi_2 : Q \to P$ which are surjective idempotent separating homomorphisms. The composition $\psi_1 \circ \psi_2$ of two surjective idempotent separating homomorphisms is also a surjective idempotent separating homomorphism. Therefore $[P] \leq_c [R]$ and $\leq_c$ is transitive and we have shown that $\leq_c$ is a partial order. □

The embedding ordering is now shown to be a partial order.

Proposition 5.2.3. *Let S be an inverse semigroup and let X denote the collection of all E-unitary inverse covers of S, up to isomorphism. Then $\leq_e$ is a partial order.*

Proof. If $P \in X$ the identity function on P is a monomorphism therefore $P \leq_e P$ so $\leq_e$ is reflexive. Now assume $P,Q \in X$ such that $P \leq_e Q$ and $Q \leq_e P$. Then there must exist $\theta_1 : Q \to P$ and $\theta_2 : P \to Q$ which are monomorphisms. Since θ_1, θ_2 are both injective P and Q must be isomorphic. Therefore P equals Q and $\leq_e$ is anti-symmetric. Next assume that $P,Q,R \in X$ satisfy $P \leq_e Q$ and $Q \leq_e R$. Then there exists $\theta_1 : R \to Q$ and $\theta_2 : Q \to P$ which are monomorphisms. The composition $\theta_1 \circ \theta_2$ of two monomorphisms is also a monomorphism. Therefore $P \leq_e R$ and $\leq_e$ is transitive and we have shown that $\leq_e$ is a partial order. □

As we alluded to at the beginning of this section, it would be nice to find a 'best' cover for a given inverse semigroup S. The word 'best' could be seen as meaning uniquely minimal with respect to a sensible ordering on the covers of S. A partial order does not guarantee a unique minimal element. In the following section we will build examples that show the relations we have just defined need not have a unique minimal element.

5.3 E-unitary covers for Clifford semigroups

In this section we present various examples of E-unitary covers of inverse semigroups which show that the composition order and the embedding order need not have a least element. Furthermore we argue these examples show that, in general, there is no sensible partial order on E-unitary covers of an inverse semigroup which has a least element, up to isomorphism. We choose to focus covers of Clifford semigroups which are relatively simple to describe whilst still being complex enough to provide the counter examples we seek.

Proposition 5.3.1. *A semilattice of groups $(E, \{G_e : e \in E\}, \{\phi_{e,f} : e,f \in E \text{ and } e \geq f\})$ is an E-unitary inverse semigroup if and only if for all $e,f \in E$ such that $e \geq f$ the homomorphism $\phi_{e,f}$ is injective.*

Proof. Let E be a semilattice. Let $G = \{G_e : e \in E\}$ be a set of groups. Let $\phi = \{\phi_{e,f} : G_e \to G_f \mid e,f \in E \text{ and } e \geq f\}$. Let $S = (E,G,\phi)$ be a semilattice of groups. The idempotents of S are the identities $\{1_{G_e} : e \in E\}$. Let $e,f \in E$ be arbitrary. If g is any element of G_f then

$$\begin{aligned} 1_e g \in E(S) &\iff (1_{G_e})\phi_{e,ef}(g)\phi_{f,ef} \in E(S) \\ &\iff 1_{ef}(g)\phi_{f,ef} \in E(S) \\ &\iff (g)\phi_{f,ef} = 1_{ef} \end{aligned}$$

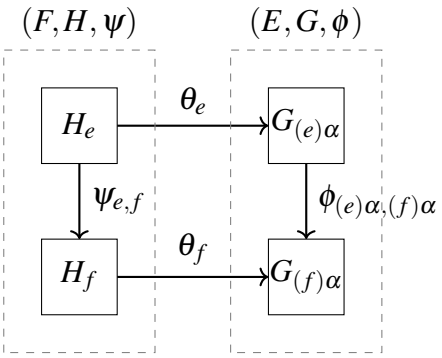


Fig. 5.4 In order for $\theta : (F,H,\psi) \to (E,G,\phi)$ to be a homomorphism this diagram must commute for all $e, f \in F$ such that $e \geq f$.

Say that there are $e, f \in E$ such that $\phi_{e,f}$ is not injective. Then there is some $g \in G_e$ where $g \neq 1_{G_e}$ and $(g)\phi_{e,f} = 1_{G_f}$. In this case we have that $1_{G_f} * g = 1_{G_f}$ even and S would not be E-unitary. It follows that S is E-unitary if and only if all homomorphisms in ϕ are injective. □

The following result shows that the possibilities for an E-unitary cover of a Clifford semigroup are fairly restricted, which makes them relatively simple examples to consider.

Proposition 5.3.2. *An E-unitary inverse semigroup which covers a semilattice of groups S must also be isomorphic to a semilattice of groups.*

Proof. This follows from the fact that idempotent separating homomorphisms preserve Green's relations. □

Now we describe the necessary and sufficient conditions for an E-unitary semilattice of groups to be a cover of a specific semilattice of groups.

Theorem 5.3.3. *Let E and F be semilattices. Let $G = \{G_e : e \in E\}$ and $H = \{H_f : f \in F\}$ be sets of groups. Let $\phi = \{\phi_{e,f} : G_e \to G_f \mid e, f \in E \text{ and } e \geq f\}$ and $\psi = \{\psi_{e,f} : H_e \to H_f \mid e, f \in F \text{ and } e \geq f\}$ be sets of group homomorphisms. Suppose that the semilattice of groups $P = (F,H,\psi)$ is an E-unitary inverse semigroup. Then there exists a surjective idempotent separating homomorphism $\theta : P \to S$ if and only if*

(i) *there exists an isomorphism $\alpha : F \to E$; and*

(ii) *there exists a set of surjective group homomorphisms $\{\theta_f : H_f \to G_{(f)\alpha} : f \in F\}$; and*

(iii) *for all $e, f \in F$ such that $e \geq f$, and for all $x \in H_e$ we have*

$$((x)\psi_{e,f})\theta_f = ((x)\theta_e)\phi_{(e)\alpha,(f)\alpha}.$$

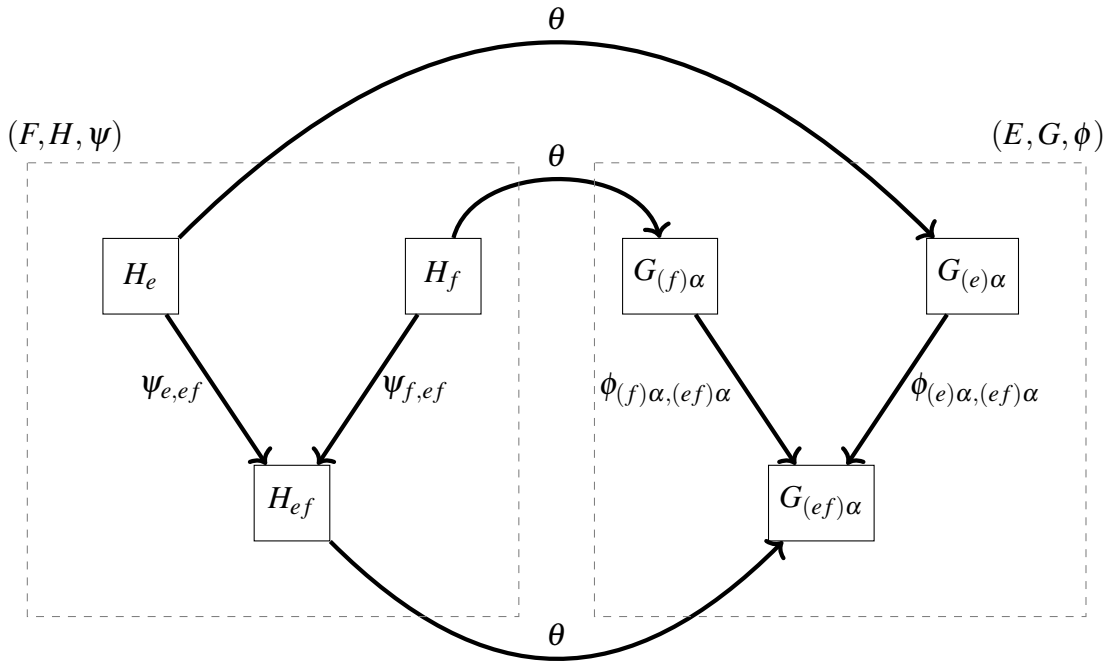


Fig. 5.5 In order for $\theta : (F,H,\psi) \to (E,G,\phi)$ to be a homomorphism this diagram must commute for all $e, f \in F$.

(Equivalently, for all $e, f \in F$ such that $e \geq f$ the diagram in Figure 5.5 must commute.)

In this case, the map $\theta : P \to S$ defined by $(x)\theta = (x)\theta_f$ where $f \in F$ satisfies $x \in H_f$ is a surjective idempotent separating homomorphism.

Proof. Assume there exists an idempotent separating homomorphism $\theta : P \to S$. Then the restriction of θ to $E(P)$ is an isomorphism from F to E and we will denote this isomorphism by α. Since θ is idempotent separating we have that $x \mathscr{H} y$ if and only if $x\theta \mathscr{H} y\theta$. It follows that for each $f \in F$ we have $(H_f)\theta \subseteq G_{(f)\alpha}$ and $(G_{(f)\alpha})\theta^{-1} \subseteq H_f$. Since θ is also surjective it must be the case that $(H_f)\theta = G_{(f)\alpha}$ for all $f \in F$, so the restriction of θ to H_f is a surjective group homomorphism. In order to show condition (iii) we use the fact that θ is a homomorphism and therefore $(xy)\theta = (x)\theta\,(y)\theta$ for all $x, y \in P$. This situation is shown in Figure 5.5. Let $x, y \in P$ be arbitrary. Let $e, f \in F$ such that $e \geq f$. Let $x \in H_e$ and let 1_{H_f} denote the identity of H_f.

$$\begin{aligned}(x\,1_{H_f})\theta &= ((x)\psi_{e,f}\,(1_{H_f})\psi_{f,f})\theta_f \\ &= ((x)\psi_{e,f})\theta_f\,(1_{H_{(f)\alpha}})\theta_f \\ &= ((x)\psi_{e,f})\theta_f\,1_{G_{(f)\alpha}} \\ &= ((x)\psi_{e,f})\theta_f\end{aligned}$$

and also

$$\begin{aligned}(x)\theta\,(1_{H_f})\theta &= (x)\theta_e\,(1_{H_f})\theta_f \\ &= ((x)\theta_e)\phi_{(e)\alpha,(f)\alpha}\,(1_{G_{(f)\alpha}})\phi_{(f)\alpha,(f)\alpha} \\ &= ((x)\theta_e)\phi_{(e)\alpha,(f)\alpha}\,1_{G_{(f)\alpha}} \\ &= ((x)\theta_e)\phi_{(e)\alpha,(f)\alpha}\end{aligned}$$

It follows that

$$(xy)\theta = (x)\theta\,(y)\theta \implies ((x)\psi_{e,ef})\theta_{ef}\,((y)\psi_{f,ef})\theta_{ef} = ((x)\theta_e)\phi_{(e)\alpha,(ef)\alpha}\,((y)\theta_f)\phi_{(f)\alpha,(ef)\alpha}$$

Therefore condition (iii) must hold if θ is a homomorphism.

To show the reverse implication we assume conditions $(i) - (iii)$ hold and construct a surjective idempotent separating homomorphism $\theta : P \to S$. Let $\theta : P \to S$ be defined by $(x)\theta = (x)\theta_f$ where $f \in F$ satisfies $x \in H_f$. If $(1_{H_e})\theta = (1_{H_f})\theta$ then $1_{G_{(e)\alpha}} = 1_{G_{(f)\alpha}}$ since α is an isomorphism and so e equals f. Therefore θ is idempotent separating. Since α is an isomorphism, for each $e \in E$ there is $H_{(e)\alpha^{-1}} \in H$ such that the image $(H_{(e)\alpha^{-1}})\theta$ is contained in G_e. Since each of the elements of $\{\theta_f : f \in F\}$ is surjective we have $(H_{(e)\alpha^{-1}})\theta = G_e$.

Therefore θ is surjective. Finally we show that θ is a homomorphism. Let $e, f \in F$ and choose $x \in H_e$ and $y \in H_f$. Since $e \geq ef$ condition (iii) tells us that

$$((x)\psi_{e,ef})\theta_{ef} = ((x)\theta_e)\phi_{(e)\alpha,(ef)\alpha},$$

and

$$((y)\psi_{f,ef})\theta_{ef} = ((y)\theta_f)\phi_{(f)\alpha,(ef)\alpha}.$$

It follows that

$$\begin{aligned}(xy)\theta &= ((x)\psi_{e,ef}\,(y)\psi_{f,ef})\theta_{ef} \\ &= ((x)\psi_{e,ef})\theta_{ef}\,((y)\psi_{f,ef})\theta_{ef} \\ &= ((x)\theta_e)\phi_{(e)\alpha,(ef)\alpha}\,((y)\theta_f)\phi_{(f)\alpha,(ef)\alpha} \\ &= (x)\theta_e\,(y)\theta_f \\ &= (x)\theta\,(y)\theta.\end{aligned}$$

Therefore θ must be a homomorphism. □

We now create special notation for semilattices of groups where the semilattice has only two elements. Let $E = \{e, f\}$ be a semilattice such that $e > f$. Let G_e and G_f be groups. Let $\phi_{e,e}$ and $\phi_{f,f}$ be the identity maps on G_e and G_f, respectively. Let $\phi_{e,f} : G_e \to G_f$ be a group homomorphism. Then $S = (E, \{G_e, G_f\}, \{\phi_{e,e}, \phi_{e,f}, \phi_{f,f}\})$ is a semilattice of groups. Herein we will refer to S by $(G_e, G_f, \phi_{e,f})$ and it will be assumed that the first group corresponds to the greater element of the semilattice. Furthermore the homomorphisms $\phi_{e,e} : G_e \to G_e$ and $\phi_{f,f} : G_f \to G_f$ which are required to define the semilattice of groups are assumed to be the identity maps on G_e and G_f, respectively. We now state a corollary to Theorem 5.3.3 which is the specialisation to the case where the size of the semilattice is two.

Corollary 5.3.4. *Let G_1, G_2, H_1 and H_2 be groups. Let $\phi : G_1 \to G_2$ and $\psi : H_1 \to H_2$ be group homomorphisms. Furthermore let ψ be injective so that $P = (H_1, H_2, \psi)$ is E-unitary. Then P is an E-unitary cover of $S = (G_1, G_2, \phi)$ if and only if there exists surjective homomorphisms $\theta_1 : H_1 \to G_1$ and $\theta_2 : H_2 \to G_2$ which satisfy:*

$$((h)\psi)\theta_2 = (h)\theta_1)\phi \text{ for all } h \in H_1$$

(or, equivalently, Figure 5.6 commutes). In this case the map $\theta : P \to S$ defined by $(h)\theta = (h)\theta_1$ if $h \in H_1$ and $(h)\theta = (h)\theta_2$ if $h \in H_2$ is a surjective idempotent separating homomorphism.

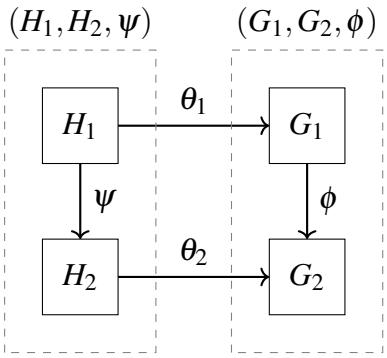


Fig. 5.6 There exists a surjective idempotent separating homomorphism from the E-unitary semilattices of group (H_1, H_2, ψ) to the semilattice of groups (G_1, G_2, ϕ) exactly when the conditions described in Corollary 5.3.4 hold.

Proof. Condition (i) of Theorem 5.3.3 always holds in this case. Condition (ii) holds exactly when there exist surjective homomorphisms $H_1 \to G_1$ and $H_2 \to G_2$. Finally, in this case there are three distinct cases for condition (iii) $e = f = 1$, $e = f = 2$, and $1 = e > f = 2$. In the cases where e and f are equal the statement follows immediately from the fact that $\psi_{1,1}, \psi_{2,2}, \phi_{1,1}$ and $\phi_{2,2}$ are the identity mappings on their respective domains. The remaining case is

$$((h)\psi)\theta_2 = ((h)\theta_1)\phi) \text{ for all } h \in H_1.$$

This proves the result.

□

Next we present our example. Let C_i denote the cyclic group of order i generated by the permutation $(1, \ldots, i)$. Let S_i denote the symmetric group of order i acting on the set $\{1, \ldots, i\}$. The example shows that neither the composition ordering nor the embedding ordering must have a least element.

Example 5.3.5. Let $\phi : C_2 \to C_3$ be the group homomorphism, where $(1,2)\phi = 1_{C_3}$. Let $\psi_1 : C_2 \to C_6$ be the group homomorphism where $(1,2)\psi_1 = (1,4)(2,5)(3,6)$. Let $\psi_2 : C_2 \to S_3$ be the group homomorphism where $(1,2)\psi_2 = (1,2)(3)$. Then the semilattice of groups $S = (C_2, C_3, \phi)$, which is not E-unitary, is covered by both the E-unitary semilattices of groups: $P_1 = (C_2, C_6, \psi_1)$ and $P_2 = (C_2, S_3, \psi_2)$. The idempotent separating homomorphism $\theta : P_1 \to S$ is formed from the isomorphism $\theta_1 : C_2 \to C_2$ and the surjective homomorphism $\theta_2 : C_6 \to C_3$ which maps $(1,2,3,4,5,6)$ to $(1,2,3)$. The idempotent separating homomorphism $\theta' : P_2 \to S$ is formed from the isomorphism $\theta'_1 : C_2 \to C_2$ and the surjective homomorphism $\theta'_2 : S_3 \to C_3$

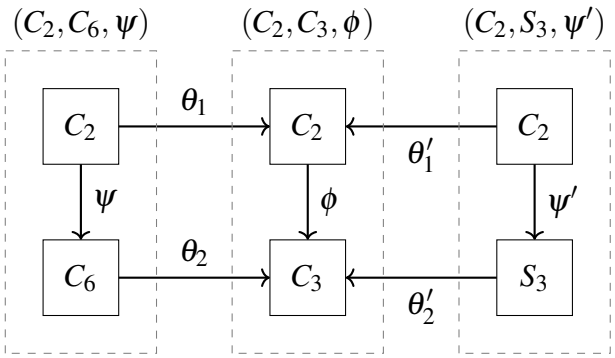


Fig. 5.7 Two non-isomorphic covers of (C_2, C_3, ϕ) of minimal order.

which maps $(1,2,3)$ to $(1,2,3)$ and maps $(1,2)$ to the identity of C_3. The covers and the surjective idempotent separating homomorphisms are illustrated in Figure 5.7.

We have provided two examples of E-unitary inverse semigroups $P_1 = (C_2, C_6, \psi_1)$ and $P_2 = (C_2, S_3, \psi_2)$ of order 8 which cover S. We will show that a cover of smaller order does not exist. First, an E-unitary cover P of S must be a semilattice of groups with semilattice of size two. Let G and H be groups. Let $\psi : G \to H$ be a homomorphism. Assume $P \cong (G, H, \psi)$. Then by Corollary 5.3.4 there must exist surjective group homomorphisms $\theta_1 : G \to C_2$ and $\theta_2 : H \to C_3$. Therefore $G/\ker_{\theta_1} \cong C_2$ and $H/\ker_{\theta_2} \cong C_3$. It follows that the order of G is at least 2 and the order of H is at least 3. Henceforth we assume $|G| = 2$ and claim that it is easy to see that any other choice would lead to a larger order for P. Since P is E-unitary, ψ must be injective and there must be a subgroup of H isomorphic to G. Therefore H has size at least 3, has a normal subgroup N such that $H/N \cong C_3$, and has a subgroup isomorphic to C_2. The least possible order for H is 6 with either $H = C_6$ or $H = S_3$. Therefore the least possible order for (G, H, ψ) is at least the sum of the least possible orders for G and H which we have shown is at least eight.

Now we can deduce that, in general, the composition ordering and the embedding ordering do not have a least element.

Theorem 5.3.6. *In general, the composition ordering $\leq_c$ and the embedding ordering $\leq_e$ on the collection (up to isomorphism) of E-unitary covers of an inverse semigroup need not have a least element.*

Proof. Let P and Q are non-isomorphic E-unitary covers of an inverse semigroup S such that either $P \leq_c Q$ or $P \leq_e Q$. Then the order of P must be less than the order of Q. Yet in Example 5.3.5 there is an inverse semigroup with two non-isomorphic E-unitary covers of the

same order. This order was shown to the minimal possible order. Thus there can be no unique minimal cover with respect to either ordering in that case. □

It seems unlikely to the author that a sensible ordering would choose either P_1 or P_2 from Example 5.3.5 as a 'superior' cover to the other. Therefore we can make no further progress without reformulating our question. If the fact that these two covers are non-isomorphic is a problem for our aim of finding a minimal cover then we might decide that we should consider the set of E-unitary covers of an inverse semigroup up to some other equivalence which is coarser than isomorphism. In the next section we give an example of such a relation for the E-unitary covers of a semilattice of groups with two idempotents.

5.4 Composition Equivalence

First, recall that the composition series of a group G is a series of finite length

$$1 = H_0 \triangleleft H_1 \triangleleft \cdots \triangleleft H_n = G,$$

with strict inclusions, such that H_i is a maximal normal subgroup of H_{i+1} for $i \in [n-1]$. The factor groups, H_{i+1}/H_i, are called composition factors. Note that the composition series of a group is not unique, and two non-isomorphic groups may both share a composition series. However, the Jordan-Hölder theorem states that any two composition series of a group will have the same length of compositions series, and the same composition factors up to permutation and isomorphism.

Now, let S be a Clifford semigroup with exactly two idempotents. Let X denote the collection of all E-unitary covers of S. We can define an equivalence relation $\sim \subset X \times X$ by $(G_1, G_2, \phi) \sim (H_1, H_2, \psi)$ if and only if there exists composition series for G_1 and G_2 which have isomorphic composition factors to the composition series for H_1 and H_2, respectively. This certainly solves our issue in Example 5.3.5 since S_3 and C_6 both have composition series with C_3 and C_2 as the only two composition factors.

Promisingly, we have tackled a class of examples which previously hindered us. Our hope is that we can prove a result of the following form.

Conjecture 5.4.1. *Let S be a Clifford semigroup with two idempotents. Let X_S be the collection of all E-unitary covers of S. Then the composition ordering $\leq_c$ is a semilattice on $X_S/\sim$.*

However we are now impeded by our knowledge of groups with identical composition factors. We will now present a sub-problem which, as far as the author is aware, we cannot make any further progress with the existing theory. First we prove the existence of a certain cover for all Clifford semigroups with two idempotents.

Proposition 5.4.2. *Let S be a Clifford semigroup with exactly two idempotents. Assume there are groups G_1, G_2 and a group homomorphism $\phi : G_1 \to G_2$ such that $S = (G_1, G_2, \phi)$. Define $\psi : G_1 \to G_1 \times G_2$ by $g\psi = (g, g\phi)$. Then the semigroup $P = (G_1, G_1 \times G_2, \psi)$ is an E-unitary cover of S.*

Proof. Let $\theta : S \to (G_1, G_1 \times G_2, \psi)$ be defined by $g_1\theta = g_1$ for $g_1 \in G_1$ and $(g_1, g_2)\theta = g_1$ for $g_1 \in G_1$. Then θ is an idempotent separating homomorphism. □

Proposition 5.4.2 leads us to pose a more specific conjecture than Conjecture 5.4.1 which despite the simplification still appears intractable to the author.

Conjecture 5.4.3. *Let $S = (G_1, G_2, \phi)$ be a Clifford semigroup with two idempotents. Let X_S^* be the collection of all E-unitary covers of S which lie below the cover $(G_1, G_1 \times G_2, \phi)$ with respect to the composition ordering on covers of S. Then the composition ordering $\leq_c$ is a semilattice on $X_S^* / \sim$.*

To see how this problem is equivalent to a problem in group theory we first make the following proposition.

Proposition 5.4.4. *Let $S = (G_1, G_2, \phi)$ be a Clifford semigroup with two idempotents. Then $(G_1, K, \psi) \in X_S$ if and only if:*

(i) *The homomorphism $\psi : G_1 \hookrightarrow K$ is injective;*

(ii) *There exists a surjective group homomorphism $\theta : K \twoheadrightarrow G_2$; and*

(iii) *For all $x \in G_1$ we have $x\phi = (x\psi)\theta$.*

It follows from Proposition 5.4.4 that elements of E-unitary covers of $S = (G_1, G_2, \phi)$ of the form (G_1, K, ψ) correspond with triples (K, ψ, θ) where: K is a group, $\psi : G_1 \to K$ is an injective homomorphism, $\theta : K \to G_2$ is a surjective homomorphism, and the diagram in Figure 5.8 commutes. We are interested in the possibilities for K. Thus we denote by $\mathscr{K}(G_1, G_2, \phi)$ the collection of groups

$$\{K : (G_1, K, \psi) \in X_{(G_1, G_2, \phi)}\}.$$

Note that $\mathscr{K}(G_1, G_2, \phi)$ is never empty. As shown in Proposition 5.4.2, we can always choose $K = G_1 \times G_2$, and define the homomorphisms ψ, θ by: $g\psi = (g, g\phi)$, and $(g_1, g_2)\theta = g_2$. Conjecture 5.4.3 corresponds to finding the 'minimal' element of $\mathscr{K}(G_1, G_2, \phi)$ which is also

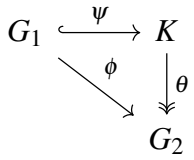


Fig. 5.8 Given G_1, G_2 and ϕ we wish to find possible K, ψ, θ such that this diagram commutes.

a homomorphic image of $G_1 \times G_2$. We will write $\mathscr{K}^*(G_1, G_2, \phi)$ to denote the subset of $\mathscr{K}(G_1, G_2, \phi)$ of groups which are homomorphic images of $G_1 \times G_2$.

The composition ordering on X_S^* corresponds to the order $\geq_h$ on $\mathscr{K}^*(G_1, G_2, \phi)$ where $K_2 \geq_h K_1$ if there exists a surjective group homomorphism $K_2 \twoheadrightarrow K_1$. Let $\approx$ denote the relation $K_1 \approx K_2$ if K_1 and K_2 have identical composition factors up to permutation and isomorphism. Then the $\approx$-classes of $\mathscr{K}^*(G_1, G_2, \phi)$ correspond to the $\sim$-classes of X_S^*. It follows that Conjecture 5.4.3 is equivalent to the following conjecture.

Conjecture 5.4.5. *Let G_1, G_2 be finite groups. Let $\phi : G_1 \to G_2$ be a homomorphism. Then the ordering $\leq_h$ is a semilattice on $\mathscr{K}^*(G_1, G_2, \phi)/\approx$.*

We end this section with an example demonstrating some thus far unseen features of this problem.

Example 5.4.6. One could suppose that any minimal $K \in \mathscr{K}^*(G_1, G_2, \phi)$ will be an extension of G_2 by $\ker\phi$. However, the author has examined the case where $G_1 = C_9$, $G_2 = S_4$, and the homomorphism ϕ is defined by $(1\,2\,3\,4\,5\,6\,7\,8\,9)\phi = (1\,2\,3))$. Using the Small Groups Library in GAP to examine all groups of order $72 = 24 * 3 = |S_4| * |\ker\phi|$ up to isomorphism, we discovered that there is no group K of order 72 such that there exists both a monomorphism $\psi : C_9 \to K$ and an epimorphism $\theta : K \to S_4$. If there were an element of $\mathscr{K}^*(G_1, G_2, \phi)$ which is an extension of S_4 by C_6 or S_3 then it would be incomparable to $S_4 \times C_9 \in \mathscr{K}^*(G_1, G_2, \phi)$. This would imply that working 'up to composition factors' is inadequate as was 'up to isomorphism'. However, by using GAP to analyse the groups of order $24 * 6$ we can say there is no such solution. Therefore the only, and thus minimal, element of $\mathscr{K}^*(G_1, G_2, \phi)$ in $S_4 \times C_9$.

The author would ideally have examined more examples (G_1, G_2, ϕ) where there is no element of $\mathscr{K}^*(G_1, G_2, \phi)$ with order equal to $|\ker\phi||G_2|$. However examining all groups of several orders for certain subgroups and certain homomorphic images quickly becomes infeasible as the orders rise. More examples may have provided valuable intuition however we were unable to find other examples of small enough order to be fully analysed in GAP.

Certainly a group $K \in \mathscr{K}^*(G_1, G_2, \phi)$ must have order at least $|G_2||\ker\phi|$ and the bound is met exactly when K is an extension of G_2 by $\ker\phi$. However Example 5.4.6 demonstrates that

there need not exist an element of this order. Of course, $G_1 \times G_2$ is an element of $\mathscr{K}^*(G_1, G_2, \phi)$. Therefore all $K \in \mathscr{K}^*(G_1, G_2, \phi)$ have order greater than or equal to $|G_2| * |\ker \phi|$ and less than or equal to $|G_1||G_2|$. It would be interesting to be able to determine the minimal order of an element of $\mathscr{K}^*(G_1, G_2, \phi)$ for any G_1, G_2, and ϕ. Moreover whether or not incomparable elements can occur remains unanswerable to us.

Chapter 6

Computing presentations of semigroups

6.1 Factorisable monoids

There are various definitions of the term factorisable for monoids. Let M be a monoid. Let A and B be subsemigroups of M. Then it is typically required that M is a monoid satisfying

$$M = AB = \{ab : a \in A, b \in B\}.$$

In some cases A, B are allowed to be subsets rather than subsemigroups, allowing more possibilities. Another variation requires that $|M| = |A||B|$ so that each ab for $a \in A$ and $b \in B$ is unique. Yet another variation could be that $A \cap B = \{1_M\}$. A *factorisable inverse monoid* is an inverse monoid M where $M = E(M)U(M)$. Factorisable groups and factorisable inverse monoids are well studied examples of factorisable monoids. We will ...

6.1.1 Factorisable orthodox monoids

We will consider the case where M is an orthodox monoid and will generalize the results of Easdown, East and FitzGerald [10] on presentations of factorisable inverse monoids. The subsemigroup of idempotents of an orthodox monoid can be any band with identity, unlike the inverse monoid case where it must be a semilattice. Let M be an orthodox monoid. Let $G = U(M)$ be the group of units of M. Let $E = E(M)$ be the band of idempotents of M. Furthermore, suppose that for each g in G there is an automorphism $\psi_g : E \to E$ defined by $e \mapsto e^g$ such that the map $\psi : G \to G$ defined by $g \mapsto \psi_g$ is an antihomomorphism. We then form the semidirect product

$$E \rtimes G = E \rtimes_\psi G = \{(e, g) | e \in E, g \in G\}$$

with multiplication defined by

$$(e_1, g_1)(e_2, g_2) = (e_1 e_2^{g_1}, g_1 g_2).$$

Let $(1, G) = \{(1, g) | g \in G\}$ and $(E, 1) = \{(e, 1) | e \in E\}$ then we have the following.

Lemma 6.1.1. *The monoid $E \rtimes G$ is a factorisable orthodox monoid with group of units $(1, G) \cong G$ and band of idempotents $(E, 1) \cong E$.*

Proof. The idempotents of $E \rtimes G$ are the set $(E, 1)$ since $(e, g) * (e, g) = (ee^g, g^2) = (e, g)$ if and only if $g = 1$ and thus the idempotents form a subsemigroup isomorphic to E so $E \rtimes G$ is orthodox. It is also clear that the identity of $E \rtimes G$ is $(1, 1)$ and that $(1, G)$ are the units. □

Again suppose that we have a factorisable orthodox monoid $M = EG$. We may define a congruence $\sim$ on $E \rtimes G$ by

$$(e, g) \sim (f, h) \text{ if and only if } e = fhg^{-1}.$$

Let G_e denote the stabilizer of e under the action $g : s \mapsto gs$ of G on M by right multiplication. For each ordered pair $e, f \in E$ let $g_{e,f}$ denote a representative element of G which satisfies $eg_{e,f} = f$, if such an element exists. In other words, for each $e \in E$, these $g_{e,f}$ are representatives of certain cosets of G_e. Then we may equally define $\sim$ by

$$(e, g) \sim (f, h) \text{ if and only if } e^G = f^G \text{ and } hg^{-1} \in G_f g_{f,e}.$$

Now if we distinguish elements $e_1, \ldots, e_k$ so that $\{e_1^G, \ldots e_k^G\}$ is the set of orbits of M containing idempotents, with respect to the action $g : e \mapsto eg$, then we may also define $\sim$ by

$$(e, g) \sim (f, h) \text{ if and only if } \exists i : e, f \in e_i^G \text{ and } hg^{-1} \in g_{e_i,f}^{-1} G_{e_i} g_{e_i,e}.$$

As before, the $g_{a,b}$ are defined so that $ag_{a,b} = b$.

Lemma 6.1.2. *The relation $\sim$ is a congruence on $E \rtimes G$.*

Proof. Let $(e_1, g_1) \sim (f_1, h_1)$ and $(e_2, g_2) \sim (h_2, f_2)$. Then $e_1 g_1 h_1^{-1} = f_1$ and $e_2 g_2 h_2^{-1} = f_2$ hold and we have:

$$\begin{aligned}(f_1,h_1)(f_2,h_2) &= (f_1 f_2^{h_1}, h_1h_2)\\ &= ((e_1g_1h_1^{-1})(e_2g_2h_2^{-1})^{h_1}, h_1h_2)\\ &= (e_1g_1h_1^{-1}h_1e_2g_2h_2^{-1}h_1^{-1}, h_1h_2)\\ &= (e_1g_1e_2g_2h_2^{-1}h_1^{-1}, h_1h_2)\\ &= (e_1e_2^{g_1}g_1g_2h_2^{-1}h_1^{-1}, h_1h_2)\\ &\sim (e_1e_2^{g_1}, g_1g_2)\\ &= (e_1,g_1)(e_2,g_2)\end{aligned}$$

as required. □

Now define $M' = (E \rtimes G)/\sim$ and for e in E and g in G, let $[e,g]$ denote the $\sim$-class of (e,g) in $E \rtimes G$.

Proposition 6.1.3. *The map $(e,g) \mapsto [e,g]$ is injective on $(1,G)$ and $(E,1)$. Thus M' is a factorisable orthodox monoid with band of idempotents $\{[e,1] : e \in E\} \cong E$ and group of units $\{[1,g] : g \in G\} \cong G$.*

Proof. Injectivity on $(E,1)$ follows from $e^1 = e$ for all e in E. Injectivity on $(1,G)$ is clear. □

Proposition 6.1.4. *Let M be a factorisable orthodox monoid with group of units G and band of idempotents E. Then $M \cong M'$.*

Proof. We show that the map $\phi : M' \to S : [e,g] \mapsto eg$ is an isomorphism. First, ϕ is surjective since if s is an element of M then $s = eg$ for some e in E and g in G so $[e,g]$ is an element of M' and $\phi([e,g]) = s$. Next, ϕ is injective since if $[e,g],[f,h]$ are elements of M' then

$$\begin{aligned}\phi([e,g]) = \phi([f,h]) &\iff eg = fh\\ &\iff egh^{-1} = f\\ &\iff (e,g) \sim (f,h)\\ &\iff [e,g] = [f,h].\end{aligned}$$

Finally, ϕ is a homomorphism since if $[e,g],[f,h]$ are elements of M' then

$$\begin{aligned}\phi([e,g][f,h]) =& \phi([ef^g, gh])\\ =& ef^g gh\\ =& egfg^{-1}gh\\ =& egfh\\ =& \phi([e,g])\phi([f,h]).\end{aligned}$$

Therefore ϕ is an isomorphism. □

6.1.2 Presentations of factorisable orthodox monoids

In this section we generalize the method of Easdown, East and FitzGerald [10] to make use of Propositions 6.1.3 and 6.1.4 to describe a presentation of an arbitrary factorisable orthodox monoid M. We require presentations of $E = E(M)$ the subsemigroup of idempotents of M and $G = G(M)$ the group of units of M.

Now suppose that M is an arbitrary factorisable orthodox monoid and suppose that $E(M)$ and $G(M)$ have presentations $\langle X_E | R_E \rangle$ and $\langle X_G | R_G \rangle$, respectively. Let $\alpha : X_E^* \to E$ and $\beta : X_G^* \to G$ be monoid epimorphisms such that $\ker\alpha = R_E^\#$ and $\ker\beta = R_G^\#$. For each $e \in E$ we choose $\hat{e}$ in $e\alpha^{-1}$ and for each $g \in G$ we choose $\hat{g}$ in $g\beta^{-1}$. We suppose that we choose $\widehat{x\alpha} = x$ and $\widehat{y\beta} = y$ for all $x \in X_E$ and $y \in X_G$. Put

$$R_{\rtimes} = \{(yx, \widehat{x\alpha^{y\beta}}y) | x \in X_E, y \in X_G\}.$$

It is known (see reference in [10]) that $E \rtimes G$ has presentation

$$\langle X_G \cup X_E | R_G \cup R_E \cup R_{\rtimes} \rangle.$$

Suppose that for each e in E we have a generating subset Σ_e of G_e, as a monoid. Furthermore suppose that there are k orbits $o_1, \ldots, o_k$ of M containing elements of E, with respect to the action of G by right multiplication. Let $e_1, \ldots, e_k \in E$ be such that $o_i = e_i^G$ and for each $1 \leq i \leq k$ and for each $f \in e_i^G \cap E$ let $g_{e_i,f}$ be an element of G such that $eg_{e_i,f} = f$. Now we define

$$R_\sim = \{(\hat{e}_i\hat{g}, \hat{e}_i) | 1 \leq i \leq k, g \in \Sigma_e\} \quad \cup \bigcup_{1\leq i\leq k} \{(\hat{e}_i\widehat{g_{e_i,f}}, \hat{f}) | f \in e_i^G \cap E\}.$$

Theorem 6.1.5. *The factorisable orthodox monoid $M \cong (E \rtimes G)/\sim$ has presentation*

$$\langle X_G \cup X_E | R_G \cup R_E \cup R_{\rtimes} \cup R_\sim \rangle$$

Proof. Put $\approx = (R_G \cup R_E \cup R_{\rtimes} \cup R_{\sim})^{\#}$ and define $\phi : (X_G \cup X_E)^* \to (E \rtimes G)/\sim$ by

$$x\phi = [x\alpha, 1] \text{ and } y\phi = [1, y\beta] \quad \text{for all } x \in X_E, y \in X_G.$$

Then ϕ is surjective since α and β are surjective and $(E \rtimes G)/\sim$ is factorisable (so every element may be written $s = eg$). It remains to show that $\ker\phi = \approx$. Now $\approx \subseteq \ker\phi$ since the relations hold as equations in $(E \rtimes G)/\sim$ after substituting the images of the generators. Suppose $u, v \in (X_G \cup X_E)^*$ and $u\phi = v\phi$. Using $R_{\rtimes}$ to sort elements of X_G to the right, $u = \hat{e}\hat{g}$ and $v = \hat{f}\hat{h}$ for some $e, f \in E$ and some $g, h \in G$. There is an idempotent e_i of e^G such that there are relations of the form $(\hat{e}_i \hat{g_{e_i,\varepsilon}}, \hat{e_j})$ in $R_{\sim}$ for each $\varepsilon \in e^G$ and similarly an element e_j of f^G with these types of relations. But then

$$[e, g] = u\phi = v\phi = [f, h]$$

so that $e^G = f^G$, $e_i = e_j$ and $hg^{-1} \in g_{e_i,f}^{-1} G_{e_i} g_{e_i,e}$. Then we can write $hg^{-1} = g_{e_i,f}^{-1} h_1 \cdots h_k g_{e_i,e}$ for some $h_1, \ldots h_k$ in Σ_{e_i}. We can then show

$$\begin{aligned} v \approx fh &\approx e_i g_{e_i,f} h \\ &\approx e_i g_{e_i,f} h g^{-1} g \\ &\approx e_i g_{e_i,f} g_{e_i,f}^{-1} h_1 \cdots h_k g_{e_i,e} g \\ &\approx e_i h_1 \cdots h_k g_{e_i,e} g \\ &\approx e_i g_{e_i,e} g \\ &\approx eg \approx u \end{aligned}$$

$$\begin{aligned} v \approx fh &\approx e_i g_{e_i,f} h && \text{by } R_{\rtimes} \\ &\approx e_i g_{e_i,f} h g^{-1} g && g^{-1}g = 1_M \\ &\approx e_i g_{e_i,f} g_{e_i,f}^{-1} h_1 \cdots h_k g_{e_i,e} g && hg^{-1} \in g_{e_i,f}^{-1} G_{e_i} g_{e_i,e} \\ &\approx e_i h_1 \cdots h_k g_{e_i,e} && h_1, \ldots, h_k \in G_{e_i} \\ \text{by } &\approx e_i g_{e_i,e} g && by R_{\rtimes} \\ &\approx eg \approx u && \text{by } R_{\rtimes} \end{aligned}$$

as required. □

References

[1] Annin, S., Jansen, T., and SMITH, C. (2009). On k'th roots in the symmetric and alternating groups. *Pi Mu Epsilon Journal*, 12(10):581–589.

[2] Besche, H. U., Eick, B., and O'BRIEN, E. A. (2002). A millennium project: constructing small groups. *International Journal of Algebra and Computation*, 12(05):623–644.

[3] Bóna, M., McLennan, A., and White, D. (2000). Permutations with roots. *Random Structures & Algorithms*, 17(2):157–167.

[4] Cameron, P., Prellberg, T., and Stark, D. (2005). Asymptotics for incidence matrix classes.

[5] Cameron, P. J. (2017). *Notes on counting: An introduction to enumerative combinatorics*, volume 26. Cambridge University Press.

[6] Cannon, J. J. and Holt, D. F. (2006). Computing conjugacy class representatives in permutation groups. *Journal of Algebra*, 300(1):213–222.

[7] Cayley, A. (1854). On the theory of groups as depending on the symbolic equation θn= 1 phil.

[8] Distler, A. (2010). *Classification and enumeration of finite semigroups*. PhD thesis, University of St Andrews.

[9] Distler, A., Jefferson, C., Kelsey, T., and Kotthoff, L. (2012). The semigroups of order 10. In *Principles and Practice of Constraint Programming*, pages 883–899. Springer.

[10] Easdown, D., East, J., and FitzGerald, D. G. (2005). Presentations of factorizable inverse monoids. *Acta Universitatis Szegediensis, Acta Scientiarum Mathematicarum*, 71(3-4):509–520.

[11] Gorenstein, D., Lyons, R., and Solomon, R. (2018). *The Classification of the Finite Simple Groups, Number 8*. American Mathematical Soc.

[12] Green, J. A. (1951). On the structure of semigroups. *Annals of Mathematics*, 54(1):163–172.

[13] Grillet, P. A. (1996). Computing finite commutative semigroups. In *Semigroup Forum*, volume 53, pages 140–154. Springer.

[14] Harary, F., March, L., and Robinson, R. (1978). On enumerating certain design problems in terms of bicoloured graphs with no isolates. *Environment and Planning B: Planning and Design*, 5(1):31–43.

[15] Harary, F. and Palmer, E. (2014). *Graphical Enumeration*. Elsevier Science.

[16] Harrison, M. A. (1973). On the number of classes of binary matrices. *IEEE Trans. Comput.*, 22(12):1048–1052.

[17] Houghton, C. (1978a). Counting completely 0-simple and completely simple semigroups. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 79(3-4):293–297.

[18] Houghton, C. (1978b). Finite zero-simple semigroups over an elementary abelian group. *Bulletin of the Australian Mathematical Society*, 18(2):211–220.

[19] Howie, J. M. (1995). *Fundamentals of semigroup theory*, volume 12. Clarendon Oxford.

[20] Howie, J. M. (2002). Semigroups, past, present and future. In *Proceedings of the International Conference on Algebra and its Applications*, pages 6–20.

[21] (https://mathoverflow.net/users/35840/derek holt), D. H. Non-isomorphic finite simple groups. MathOverflow. URL:https://mathoverflow.net/q/107660 (version: 2012-09-20).

[22] (https://math.stackexchange.com/users/1277/yuval filmus), Y. F. Number of connected graphs on labeled vertices, counted according to parity. Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/68457 (version: 2011-09-29).

[23] (https://math.stackexchange.com/users/510535/dan carmon), D. C. Number of connected graphs on labeled vertices, counted according to parity. Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/2552887 (version: 2017-12-05).

[24] James, G. and Kerber, A. (1981). The representation theory of the symmetric group, volume 16 of encyclopedia of mathematics and its applications.

[25] Jefferson, C., Jonauskyte, E., Pfeiffer, M., and Waldecker, R. (2018). Minimal and canonical images. *Journal of Algebra*.

[26] Jovovic, V. Number of binary matrices up to row and column permutations. The On-Line Encyclopedia of Integer Sequences. URL:https://oeis.org/A005748/a005748.pdf).

[27] Jovovic, V. Sequence a028657. The On-Line Encyclopedia of Integer Sequences. URL:http://oeis.org/A028657.

[28] Kilibarda, G. and Jovović, V. (2014). Enumeration of certain classes of t_0-hypergraphs. *arXiv preprint arXiv:1411.4187*.

[29] Lawson, M. V. (1998). *Inverse semigroups: the theory of partial symmetries*. World Scientific.

[30] Leanos, J., Moreno, R., and Rivera-Martínez, L. M. (2012). On the number of mth roots of permutations. *Australas. J Comb.*, 52:41–54.

[31] Malandro, M. E. (2019). Enumeration of finite inverse semigroups. In *Semigroup Forum*, volume 99, pages 679–723. Springer.

[32] McAlister, D. B. (1974). Groups, semilattices and inverse semigroups. ii. *Transactions of the American Mathematical Society*, 196:351–370.

[33] Míšek, B. (1964). O počtu tříd silně ekvivalentních incidenčních matic. *Časopis pro pěstování matematiky [On the number of classes of strongly equivalent incidence matrices]*, 089(2):211–218.

[34] Pouyanne, N. (2002). On the number of permutations admitting an m-th root. *the electronic journal of combinatorics*, pages R3–R3.

[35] Preston, G. B. (1954). Inverse semi-groups. *Journal of the London Mathematical Society*, 1(4):396–403.

[36] Rees, D. (1940). On semi-groups. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 36, pages 387–400. Cambridge University Press.

[37] Rose, H. E. (1995). *A course in number theory*. Oxford University Press.

[38] Schottenfels, I. M. (1899). Two non-isomorphic simple groups of the same order 20, 160. *The Annals of Mathematics*, 1(1/4):147–152.

[39] Sloane, N. J. A. (2010). Sequence a181230. The On-Line Encyclopedia of Integer Sequences. URL:http://oeis.org/A181230.

[40] Vagner, V. (1952). Generalized groups. In *Dokl. Akad. Nauk SSSR*, volume 84, pages 1119–1122.

Appendix A

Details of implementation

As noted in Chapters 2, 3, and 4, several of the results in this thesis have been coded in **GAP** by the author. The **GAP** code written by the author can be found in the following GitHub repository: https://github.com/ChristopherRussell/0-simple-semigroups. This is exactly the code used to produce the results displayed in Section 2.7 and Section 4.5, as well as the code required to generate the database described in Chapter 3.

It is important to test the validity of the results returned by algorithms. The author was able to count the number of semigroups returned by the database implementation to obtain counts of the number of isomorphism classes of 0-simple semigroups and congruence free semigroups of orders at most 49. The methods required to create the database, and thus also the code, were completely different to the counting methods used to enumerate 0-simple semigroups or congruence free semigroups. Thus by seeing that the counts obtained from the database code matched those from the enumeration methods we were able to reasonably validate both algorithms. For the database code, the author tested the semigroups it returned in **GAP** using methods from the **Semigroups** library. For example we tested whether the semigroup were 0-simple semigroups, as well as whether the semigroups generated were unique up to isomorphism.

Ultimately the author intends the code to be integrated as a **GAP** package. As of the time when this thesis was completed, the code lacks the proper test coverage and documentation to be accepted as a **GAP** package. The author intends to resolve this in the near future.

Index